

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2005, 12:09:37 ; Search time 161 Seconds

(without alignments)
60.056 Million cell updates/sec

Title: US-10-694-847-3

Perfect score: 126

Sequence: 1 KISVSYDNFALVDYLVFERRKSDPTD 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 100 summaries

Database : A_GenSeq_16Dec04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	126	100.0	25	2 AAW90170	AAW90170 Triabin/c
2	119	94.4	25	2 AAW90171	AAW90171 Triabin/c
3	104.5	82.9	142	2 AARS9717	AARS9717 Thrombin
4	104.5	82.9	142	2 AARS9715	AARS9715 Thrombin
5	104.5	82.9	142	2 AARS9716	AARS9716 Thrombin
6	104.5	82.9	160	2 AARS9721	AARS9721 Thrombin
7	104.5	82.9	160	2 AARS9720	AARS9720 Thrombin
8	104.5	82.9	160	2 AARS9719	AARS9719 Thrombin
9	98.5	78.2	142	2 AARS9718	AARS9718 Thrombin
10	98.5	78.2	160	2 AARS9722	AARS9722 Thrombin
11	51	40.5	544	6 AABU31592	AABU31592 Protein e
12	51	40.5	548	7 ABO61580	ABO61580 Protein e
13	50	39.7	549	6 AABU28453	AABU28453 Protein e
14	50	39.7	565	7 ABO67213	ABO67213 Klebsiell
15	48	38.1	213	3 AAB35283	AAB35283 Human pro
16	48	38.1	782	4 AAM47219	AAM47219 Human NOV
17	48	38.1	783	4 AAG65764	AAG65764 Human pro
18	48	38.1	783	4 AAB85786	AAB85786 Human kin
19	48	38.1	786	4 AAU03518	AAU03518 Human pro
20	48	38.1	790	3 AAW90878	AAW90878 Human ker
21	48	38.1	818	8 ABM84424	ABM84424 Human dia
22	48	38.1	823	3 AAW90879	AAW90879 Human ker
23	48	38.1	1078	6 ABP96069	ABP96069 Human pro
24	47.5	37.7	321	7 ADC95348	ADC95348 E. faeciu
25	47	37.3	53	8 ADL72206	ADL72206 Mouse sal

26	47	37.3	101	4 AAU20986	AAU20986 Human nov
27	47	37.3	511	7 AAU17269	AAU17269 Novel sig
28	47	37.3	511	7 ADB93977	ADB93977 Human nov
29	47	37.3	519	6 ABM70476	ABM70476 Photocorb
30	47	37.3	522	4 AAU87118	AAU87118 Novel cen
31	47	37.3	522	4 ADI54433	ADI54433 Novel hum
32	47	37.3	896	7 ADJ70005	ADJ70005 Human hea
33	47	37.3	923	4 AAE11773	AAE11773 Human kin
34	47	37.3	926	4 AAB65631	AAB65631 Novel pro
35	47	37.3	926	8 ADI29238	ADI29238 Human MAR
36	47	37.3	931	8 ADL72200	ADL72200 Mouse sal
37	47	37.3	950	7 ADI14168	ADI14168 Human sic
38	47	37.3	958	4 AABU2990	AABU2990 Human sig
39	46.5	36.9	760	2 AAW29490	AAW29490 Programme
40	46.5	36.9	760	5 AAU75570	AAU75570 Programme
41	46	36.5	408	5 AAB50115	AAB50115 Listeria
42	46	36.5	529	5 AAU75255	AAU75255 Escherich
43	46	36.5	548	4 AAU74630	AAU74630 E. coli c
44	46	36.5	548	6 AABU28692	ABU28692 Protein e
45	46	36.5	744	4 AABG25898	ABG25898 Novel hum
46	46	36.5	954	4 AABG15950	ABG15950 Novel hum
47	46	36.5	954	4 AABG30332	ABG30332 Novel hum
48	46	36.5	954	4 AABG15937	ABG15937 Novel hum
49	46	36.5	954	4 AABG28932	ABG28932 Novel hum
50	45	35.7	776	7 ADC34766	ADC34766 Rat salt
51	45	35.7	1116	8 ADNI19416	ADNI19416 Bacterial
52	44.5	35.3	171	5 ABP30507	ABP30507 Streptococ
53	44.5	35.3	180	5 ABP28869	ABP28869 Streptococ
54	44.5	35.3	477	7 ADH87668	ADH87668 Enterococ
55	44.5	35.3	1249	6 ABU48355	ABU48355 Protein e
56	44.5	35.3	1744	4 AAB862890	AB862890 Drosophil
57	44.5	35.3	4092	8 ADS43848	ADS43848 Bacterial
58	44	34.9	177	4 AAB69506	AB69506 Drosophil
59	44	34.9	224	4 AAG65845	AAG65845 Human pho
60	44	34.9	426	4 AAB62957	AB62957 Drosophil
61	44	34.9	653	8 ADJ48801	ADJ48801 Oll-assoc
62	44	34.9	698	8 ADJ48690	ADJ48690 Oll-assoc
63	44	34.9	912	8 ADJ48689	ADJ48689 Oll-assoc
64	44	34.9	148	8 ADJ75039	ADJ75039 Murine im
65	43.5	34.5	349	6 ABU44349	ABU44349 Protein e
66	43	34.1	89	7 ADH85636	ADH85636 Enterococ
67	43	34.1	127	3 AAB41228	AAB41228 Human ORF
68	43	34.1	222	2 AAW32220	AAW32220 Bovine al
69	43	34.1	222	5 AAE17468	AAE17468 Bovine al
70	43	34.1	222	8 ADJ75378	ADJ75378 Bovine al
71	43	34.1	252	7 ADJ55449	ADJ55449 Human nov
72	43	34.1	270	5 AAU91558	AAU91558 Amino aci
73	43	34.1	283	7 ADC95791	ADC95791 E. faeciu
74	43	34.1	293	5 AAB55570	AB55570 Lactococc
75	43	34.1	293	8 ADNI8458	ADNI8458 Bacterial
76	43	34.1	425	7 AAB86077	AB86077 Streptococ
77	43	34.1	430	2 AAW03567	AAW03567 Arabidops
78	43	34.1	430	8 ADI19010	ADI19010 Arabidops
79	43	34.1	527	7 ADC96318	ADC96318 E. faeciu
80	43	34.1	583	4 AABG30934	ABG30934 Novel hum
81	43	34.1	963	6 AABU33173	ABU33173 Protein e
82	43	34.1	1464	4 AAM41091	AAM41091 Human pol
83	43	34.1	1464	4 AAM41089	AAM41089 Human pol
84	43	34.1	1464	4 AAM41090	AAM41090 Human pol
85	42.5	33.7	105	6 AABG76179	ABG76179 Human Q9H
86	42.5	33.7	115	6 AABG76175	ABG76175 Human Q9H
87	42.5	33.7	116	6 AABG76177	ABG76177 Human Q9H
88	42.5	33.7	121	6 AABG76178	ABG76178 Human Q9H
89	42.5	33.7	121	2 AAY02285	AY02285 Secreted
90	42.5	33.7	131	6 AABG76174	ABG76174 Human Q9H
91	42.5	33.7	131	6 AABG76175	ABG76175 Human Q9H
92	42.5	33.7	131	8 ADQ75040	ADQ75040 Human imm
93	42.5	33.7	131	8 ADP25093	ADP25093 PRO polyp
94	42.5	33.7	132	6 AABG76176	ABG76176 Human sec
95	42.5	33.7	132	6 AABG76176	ABG76176 Human sec
96	42.5	33.7	139	4 AAM25958	AM25958 Human pro
97	42.5	33.7	182	4 AAB11926	AB11926 Human sec
98	42.5	33.7	183	6 ADA05772	DA05772 Human NOV

99 42.5 33.7 183 8 Adn62936 Human NOV
100 42 33.3 31 2 AAW35727 Bovine ml

ALIGNMENTS

RESULT 1

AAW90170 standard; peptide; 25 AA.

AC AAW90170;

DT 15-MAR-1999 (first entry)

DE Triabin/thrombin-binding domain cyclic peptide analogue #2.

KW Thrombin inhibitor; cyclic peptide; medicament; thrombosis; treatment;
KM unstable angina; arteriosclerosis; post-angioplasty restenosis; triabin;
myocardial infarction; thrombolysis; blood coagulation; haemodialysis.

OS Synthetic.

FT Key Location/Qualifiers

FT Modified-site 1..13
/note="the side chains of Lys(1) and Asp(13) condense to form a beta-lactam bridge"

FT Misc-difference 14
/note="D form residue"

DE19724791-A1.

PD 10-DEC-1998.

PF 06-JUN-1997; 97DE-01024791.

PR 06-JUN-1997; 97DE-01024791.

PA (SCHD) SCHERING AG.

PI Noeske-Jungblut C, Egner U, Donner P, Schleuning W, Bode W;

PI Prior PF;

DR WPI; 1999-036195/04.

PT Cyclic peptide analogues of triabin thrombin-binding domain - useful as
thrombin inhibitors.

PS Claim 7; Page 7; 10pp; German.

CC This sequence is an example of a novel cyclic peptide of formula Y1-X1-
Ser-X2-Ser-X3-X4-Asn-Phe-X5-X6-X7-Y2-D-Tyr-X8-Val-X9-Glu-X10-X11-X12-
Ser-X13-X14-Asp where one of Y1 and Y2 is Phe, Lys, Cys or Orn and the
other is Asp, Cys or Glu, and Y1 and Y2 are linked together through a
side chain or a beta-turn mimetic; X1-X14 are any amino acids and can be
linked together through side chains. Examples of such peptides are stated
to be thrombin inhibitors and their use to prepare medicaments for
treating thrombosis, unstable angina or arteriosclerosis, for preventing
post-angioplasty restenosis, for treating myocardial infarction after
thrombolysis or for inhibiting blood coagulation during haemodialysis is
claimed

CC Sequence 25 AA;

Query Match 100.0%; Score 126; DB 2; Length 25;

Best Local Similarity 100.0%; Pred. No. 3.4e-13;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 KISVSYDNFALVDYLVFERTKSDTD 25
1 KISVSYDNFALVDYLVFERTKSDTD 25

RESULT 2
AAW90171
ID AAW90171 standard; peptide; 25 AA.

AC AAW90171;

DT 15-MAR-1999 (first entry)

DE Triabin/thrombin-binding domain cyclic peptide analogue #2.

KW Thrombin inhibitor; cyclic peptide; medicament; thrombosis; treatment;
KM unstable angina; arteriosclerosis; post-angioplasty restenosis; triabin;
myocardial infarction; thrombolysis; blood coagulation; haemodialysis.

OS Synthetic.

FT Key Location/Qualifiers

FT Modified-site 1..13
/note="the side chains of Lys(1) and Asp(13) condense to form a beta-lactam bridge"

FT Modified-site 7..19
/note="the side chains of Glu(7) and Lys(19) condense to form a beta-lactam bridge"

FT Misc-difference 14
/note="D form residue"

DE19724791-A1.

PD 10-DEC-1998.

PF 06-JUN-1997; 97DE-01024791.

PR 06-JUN-1997; 97DE-01024791.

PA (SCHD) SCHERING AG.

PI Noeske-Jungblut C, Egner U, Donner P, Schleuning W, Bode W;

PI Prior PF;

DR WPI; 1999-036195/04.

PT Cyclic peptide analogues of triabin thrombin-binding domain - useful as
thrombin inhibitors.

PS Claim 7; Page 8; 10pp; German.

CC This sequence is an example of a novel cyclic peptide of formula Y1-X1-
Ser-X2-Ser-X3-X4-Asn-Phe-X5-X6-X7-Y2-D-Tyr-X8-Val-X9-Glu-X10-X11-X12-
Ser-X13-X14-Asp where one of Y1 and Y2 is Phe, Lys, Cys or Orn and the
other is Asp, Cys or Glu, and Y1 and Y2 are linked together through a
side chain or a beta-turn mimetic; X1-X14 are any amino acids and can be
linked together through side chains. Examples of such peptides are stated
to be thrombin inhibitors and their use to prepare medicaments for
treating thrombosis, unstable angina or arteriosclerosis, for preventing
post-angioplasty restenosis, for treating myocardial infarction after
thrombolysis or for inhibiting blood coagulation during haemodialysis is
claimed

CC Sequence 25 AA;

Query Match 94.4%; Score 119; DB 2; Length 25;

Best Local Similarity 92.0%; Pred. No. 4.7e-12;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 KISVSYDNFALVDYLVFERTKSDTD 25
1 KISVSYDNFALVDYLVFERTKSDTD 25

RESULT 3

AAW9717
ID AAW9717 standard; protein; 142 AA.

AC AAR59717;
 XX
 DT 25-MAR-2003 (revised)
 DT 02-MAR-1995 (first entry)
 XX
 DE Thrombin inhibitory protein T145.
 XX
 KM Thrombin: inhibition; T112; T128; T145; T15; saliva;
 KM blood sucking insect; Triatoma pallidipennis; thrombosis;
 KM arteriosclerosis; unstable angina; vascular reocclusion; coagulation;
 KM haemo-dialysis; serine protease.
 XX
 OS Triatoma pallidipennis.
 XX
 PN WO9413807-A1.
 XX
 PD 23-JUN-1994.
 XX
 PF 03-DEC-1993; 93WO-DE001172.
 XX
 PR 04-DEC-1992; 92DE-04241659.
 PR 12-FEB-1993; 93DE-04304731.
 PR 17-AUG-1993; 93DE-04328336.
 PR 25-NOV-1993; 93DE-04340798.
 XX
 XX (SCHD) SCHERING AG.
 PA Noeskejunghut C, Schlenning W, Alagon A, Possani L;
 PI Cuevas-Aguirre D, Donner P, Haendler B, Hechler U;
 PI N-PSDB; AA068270.
 DR WPI; 1994-217883/26.
 XX
 PT New thrombin inhibitory proteins from saliva of blood sucking insects -
 PT esp. Triatoma pallidipennis, for treating or preventing thrombosis,
 PT unstable angina, vascular reocclusion etc., also corresponding DNA,
 PT vectors and transformed cells.
 XX
 PS Claim 4; Page 29; 57pp; German.
 XX
 CC Thrombin inhibitory proteins are isolated from Triatoma pallidipennis.
 CC They are used to treat thrombosis, unstable angina or arteriosclerosis;
 CC to prevent vascular reocclusion after balloon angioplasty; and to inhibit
 CC coagulation during haemo-dialysis. The DNA sequences of thrombin
 CC inhibitory proteins T112, T128, T145 and T15 are given in AA068272-75
 CC (sequences encoding the corresp. mature proteins are given in AA068268-
 CC 71). The proteins are specific for thrombin, even at a 40-fold excess
 CC other serine proteases are not inhibited. (Updated on 25-MAR-2003 to
 CC correct PN field.) (Updated on 25-MAR-2003 to correct PI field.)
 CC
 XX
 SQ Sequence 142 AA;
 Query Match 82.9%; Score 104.5; DB 2; Length 142;
 Best Local Similarity 64.9%; Pred. No. 8.6e-09;
 Matches 24; Conservative 0; Mismatches 0; Indels 13; Gaps 1;
 QY 2 ISVSYNFALV-----DYLVFERTKSDTD 25
 DB 99 ISVSYNFALVCRSIIFTSQPKEDYLVFERTKSDTD 135
 RESULT 4
 AAR59715 standard; protein; 142 AA.
 AC AAR59715;
 XX
 DT 25-MAR-2003 (revised)
 DT 02-MAR-1995 (first entry)
 XX
 DE Thrombin inhibitory protein T112.
 DE Thrombin: inhibition; T112; T128; T145; T15; saliva;
 KM Thrombin; inhibition; T112; T128; T145; T15; saliva;

KM blood sucking insect; Triatoma pallidipennis; thrombosis;
 KM arteriosclerosis; unstable angina; vascular reocclusion; coagulation;
 KM haemo-dialysis; serine protease.
 XX
 OS Triatoma pallidipennis.
 XX
 PN WO9413807-A1.
 XX
 PD 23-JUN-1994.
 XX
 PF 03-DEC-1993; 93WO-DE001172.
 XX
 PR 04-DEC-1992; 92DE-04241659.
 PR 12-FEB-1993; 93DE-04304731.
 PR 17-AUG-1993; 93DE-04328336.
 PR 25-NOV-1993; 93DE-04340798.
 XX
 XX (SCHD) SCHERING AG.
 PA Noeskejunghut C, Schlenning W, Alagon A, Possani L;
 PI Cuevas-Aguirre D, Donner P, Haendler B, Hechler U;
 PI N-PSDB; AA068268.
 DR WPI; 1994-217883/26.
 XX
 PT New thrombin inhibitory proteins from saliva of blood sucking insects -
 PT esp. Triatoma pallidipennis, for treating or preventing thrombosis,
 PT unstable angina, vascular reocclusion etc., also corresponding DNA,
 PT vectors and transformed cells.
 XX
 PS Claim 4; Page 27; 57pp; German.
 XX
 CC Thrombin inhibitory proteins are isolated from Triatoma pallidipennis.
 CC They are used to treat thrombosis, unstable angina or arteriosclerosis;
 CC to prevent vascular reocclusion after balloon angioplasty; and to inhibit
 CC coagulation during haemo-dialysis. The DNA sequences of thrombin
 CC inhibitory proteins T112, T128, T145 and T15 are given in AA068272-75
 CC (sequences encoding the corresp. mature proteins are given in AA068268-
 CC 71). The proteins are specific for thrombin, even at a 40-fold excess
 CC other serine proteases are not inhibited. (Updated on 25-MAR-2003 to
 CC correct PN field.) (Updated on 25-MAR-2003 to correct PI field.)
 CC
 XX
 SQ Sequence 142 AA;
 Query Match 82.9%; Score 104.5; DB 2; Length 142;
 Best Local Similarity 64.9%; Pred. No. 8.6e-09;
 Matches 24; Conservative 0; Mismatches 0; Indels 13; Gaps 1;
 QY 2 ISVSYNFALV-----DYLVFERTKSDTD 25
 DB 99 ISVSYNFALVCRSIIFTSQPKEDYLVFERTKSDTD 135
 RESULT 5
 AAR59716 standard; protein; 142 AA.
 AC AAR59716;
 XX
 DT 25-MAR-2003 (revised)
 DT 02-MAR-1995 (first entry)
 XX
 DE Thrombin inhibitory protein T128.
 DE Thrombin: inhibition; T112; T128; T145; T15; saliva;
 KM blood sucking insect; Triatoma pallidipennis; thrombosis;
 KM arteriosclerosis; unstable angina; vascular reocclusion; coagulation;
 KM haemo-dialysis; serine protease.
 XX
 OS Triatoma pallidipennis.
 XX
 PN WO9413807-A1.

PD 23-JUN-1994.
 XX
 XX 03-DEC-1993; 93WO-DE001172.
 XX
 PR 04-DEC-1992; 92DE-04241659.
 PR 12-FEB-1993; 93DE-04304731.
 PR 17-AUG-1993; 93DE-04328336.
 PR 25-NOV-1993; 93DE-04340798.
 XX
 PA (SCHD) SCHERING AG.
 XX
 PI Noeskejunghut C, Schleunig W, Alagon A, Possant L;
 PI Cuevas-Aguirre D, Donner P, Haendler B, Hechler U;
 XX
 DR WPI; 1994-217883/26.
 DR N-PSDB; AAO68269.
 XX
 PT New thrombin inhibitory proteins from saliva of blood sucking insects -
 PT esp. Triatoma pallidipennis, for treating or preventing thrombosis,
 PT unstable angina, vascular reocclusion etc., also corresponding DNA,
 PT vectors and transformed cells.
 XX
 PS Claim 4; Page 28; 57pp; German.
 XX
 CC Thrombin inhibitory proteins are isolated from Triatoma pallidipennis.
 CC They are used to treat thrombosis, unstable angina or arteriosclerosis;
 CC to prevent vascular reocclusion after balloon angioplasty; and to inhibit
 CC coagulation during haemo-dialysis. The DNA sequences of thrombin
 CC inhibitory proteins T12, T128, T145 and T15 are given in AAO68272-75
 CC (sequences encoding the corresp. mature proteins are given in AAO68268-
 CC 71). The proteins are specific for thrombin, even at a 40-fold excess
 CC other serine proteases are not inhibited. (Updated on 25-MAR-2003 to
 CC correct PN field.) (Updated on 25-MAR-2003 to correct PI field.)
 XX
 SQ Sequence 142 AA;
 Query Match 82.9%; Score 104.5; DB 2; Length 142;
 Best Local Similarity 64.9%; Pred. No. 8.6e-09;
 Matches 24; Conservative 0; Mismatches 0; Indels 13; Gaps 1;
 Qy 2 ISVSYNFALV-----DYLVFERTSPTD 25
 Db 99 ISVSYNFALVCRSITFTSQPKEDYLVFERTSPTD 135
 RESULT 6
 AAR59721
 ID AAR59721 standard; protein; 160 AA.
 XX
 AC AAR59721;
 XX
 DT 25-MAR-2003 (revised)
 DT 02-MAR-1995 (first entry)
 XX
 DE Thrombin inhibitory protein T145.
 XX
 KW Thrombin; inhibition; T12; T128; T145; T15; saliva;
 KW blood sucking insect; Triatoma pallidipennis; thrombosis;
 KW arteriosclerosis; unstable angina; vascular reocclusion; coagulation;
 KW haemo-dialysis; serine protease.
 XX
 OS Triatoma pallidipennis.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..18
 FT /label= sig_peptide
 FT Protein 19..160
 FT /label= mat_protein
 XX
 PN MO9413807-A1.
 XX
 PD 23-JUN-1994.
 XX

PF 03-DEC-1993; 93WO-DE001172.
 XX
 XX 04-DEC-1992; 92DE-04241659.
 PR 12-FEB-1993; 93DE-04304731.
 PR 17-AUG-1993; 93DE-04328336.
 PR 25-NOV-1993; 93DE-04340798.
 XX
 PA (SCHD) SCHERING AG.
 XX
 PI Noeskejunghut C, Schleunig W, Alagon A, Possant L;
 PI Cuevas-Aguirre D, Donner P, Haendler B, Hechler U;
 XX
 DR WPI; 1994-217883/26.
 DR N-PSDB; AAO68274.
 XX
 PT New thrombin inhibitory proteins from saliva of blood sucking insects -
 PT esp. Triatoma pallidipennis, for treating or preventing thrombosis,
 PT unstable angina, vascular reocclusion etc., also corresponding DNA,
 PT vectors and transformed cells.
 XX
 PS Claim 5; Page 33; 57pp; German.
 XX
 CC Thrombin inhibitory proteins are isolated from Triatoma pallidipennis.
 CC They are used to treat thrombosis, unstable angina or arteriosclerosis;
 CC to prevent vascular reocclusion after balloon angioplasty; and to inhibit
 CC coagulation during haemo-dialysis. The DNA sequences of thrombin
 CC inhibitory proteins T12, T128, T145 and T15 are given in AAO68272-75
 CC (sequences encoding the corresp. mature proteins are given in AAO68268-
 CC 71). The proteins are specific for thrombin, even at a 40-fold excess
 CC other serine proteases are not inhibited. (Updated on 25-MAR-2003 to
 CC correct PN field.) (Updated on 25-MAR-2003 to correct PI field.)
 XX
 SQ Sequence 160 AA;
 Query Match 82.9%; Score 104.5; DB 2; Length 160;
 Best Local Similarity 64.9%; Pred. No. 9.9e-09;
 Matches 24; Conservative 0; Mismatches 0; Indels 13; Gaps 1;
 Qy 2 ISVSYNFALV-----DYLVFERTSPTD 25
 Db 117 ISVSYNFALVCRSITFTSQPKEDYLVFERTSPTD 153
 RESULT 7
 AAR59720
 ID AAR59720 standard; protein; 160 AA.
 XX
 AC AAR59720;
 XX
 DT 25-MAR-2003 (revised)
 DT 02-MAR-1995 (first entry)
 XX
 DE Thrombin inhibitory protein T128.
 XX
 KW Thrombin; inhibition; T12; T128; T145; T15; saliva;
 KW blood sucking insect; Triatoma pallidipennis; thrombosis;
 KW arteriosclerosis; unstable angina; vascular reocclusion; coagulation;
 KW haemo-dialysis; serine protease.
 XX
 OS Triatoma pallidipennis.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..18
 FT /label= sig_peptide
 FT Protein 19..160
 FT /label= mat_protein
 XX
 PN MO9413807-A1.
 XX
 PD 23-JUN-1994.
 PF 03-DEC-1993; 93WO-DE001172.
 XX

PR	04-DEC-1992;	93DE-04241659.
PR	12-FEB-1993;	93DE-04304731.
PR	17-AUG-1993;	93DE-04328336.
PR	25-NOV-1993;	93DE-04340798.
XX	(SCHD) SCHERING AG.	
XX	Noeskejungblut C, Schleuning W, Alagon A, Poesani L,	
PI	Cuevas-Aguilera D, Donner F, Haendler B, Hechler U;	
XX	WP1; 1994-217883/26.	
DR	N-PSDB; AAQ68273.	
PT	New thrombin inhibitory proteins from saliva of blood sucking insects -	
PT	esp. Triatoma pallidipennis, for treating or preventing thrombosis,	
PT	unstable angina, vascular reocclusion etc., also corresponding DNA,	
PT	vectors and transformed cells.	
XX		
PS	Claim 5; Page 32; 57pp; German.	
XX		
CC	Thrombin inhibitory proteins are isolated from Triatoma pallidipennis.	
CC	They are used to treat thrombosis, unstable angina or arteriosclerosis;	
CC	to prevent vascular reocclusion after balloon angioplasty; and to inhibit	
CC	coagulation during haemo-dialysis. The DNA sequences of thrombin	
CC	inhibitory proteins T112, T128, T145 and T15 are given in AAQ68272-75	
CC	(sequences encoding the corresp. mature proteins are given in AAQ68268-	
CC	71). The proteins are specific for thrombin, even at a 40-fold excess	
CC	other serine proteases are not inhibited. (Updated on 25-MAR-2003 to	
CC	correct PW field.) (Updated on 25-MAR-2003 to correct PI field.)	
XX		
SQ	Sequence 160 AA:	
Query Match	82.9%; Score 104.5; DB 2; Length 160;	
Best Local Similarity	64.9%; Pred. No. 9.9e-09;	
Matches 24; Conservative	0; Mismatches 0; Indels 13; Gaps 1,	
OY	2 ISVSYNPALV-----DYLVPEKTSPTD 25	
	117 ISVSYNPALVCRSIIFTSQPKEDYLVPEKTSPTD 153	
Db		
RESULT 8		
AAAS9719		
ID	AAAS9719 standard; protein; 160 AA.	
XX		
AC	AAAS9719;	
XX		
DT	25-MAR-2003 (revised)	
DT	02-MAR-1995 (first entry)	
XX		
DE	Thrombin inhibitory protein T112.	
XX		
KW	Thrombin; inhibition; T112; T128; T145; T15; saliva;	
KW	blood, sucking insect; Triatoma pallidipennis; thrombosis;	
KW	arteriosclerosis; unstable angina; vascular reocclusion; coagulation;	
KW	haemo-dialysis; serine protease.	
XX		
OS	Triatoma pallidipennis.	
XX		
FH	Key	Location/Qualifiers
FT	Peptide	1..18
FT		/label= sig_peptide
FT	Protein	19..160
FT		/label= mat_protein
XX		
PN	W09413807-A1.	
XX		
XX		
PD	23-JUN-1994.	
PF	03-DEC-1993;	93WO-DE001172.
XX		
DR	04-DEC-1992;	92DE-04241659.
DR	12-FEB-1993;	93DE-04304731.
PR		

```

PR 17-AUG-1993; 93DE-04328336.
PR 25-NOV-1993; 93DE-04340798.
PA (SCHD ) SCHERING AG.
XX
XX Noeskejungblut C, Schleuning W, Alagon A, Possani L;
PI Cuevas-Aguirre D, Donner P, Haendler B, Hechler U;
XX
XX WPI; 1994-217883/26.
DR N-PSDB; AAQ68272.
XX
XX New thrombin inhibitory proteins from saliva of blood sucking insects -
PT esp. Triatoma pallidipennis, for treating or preventing thrombosis,
PT unstable angina, vascular reocclusion etc., also corresponding DNA,
PT vectors and transformed cells.
XX
XX Claim 5; Page 31; 57pp; German.
XX
XX Thrombin inhibitory proteins are isolated from Triatoma pallidipennis.
CC They are used to treat thrombosis, unstable angina or arteriosclerosis;
CC to prevent vascular reocclusion after balloon angioplasty; and to inhibit
CC coagulation during haemo-dialysis. The DNA sequences of thrombin
CC inhibitory proteins T112, T128, T145 and T15 are given in AAQ68272-75
CC (sequences encoding the corresp. mature proteins are given in AAQ68268-
CC 71). The proteins are specific for thrombin, even at a 40-fold excess
CC other serine proteases are not inhibited. (Updated on 25-MAR-2003 to
CC correct PI field.) (Updated on 25-MAR-2003 to correct PI field.)
XX
XX Sequence 160 AA;
SQ
Query Match 82.9%; Score 104.5; DB 2; Length 160;
Best Local Similarity 64.9%; Pred. No. 9.9e-09;
Matches 24; Conservative 0; Mismatches 0; Indels 13; Gaps 1;
Oy 2 ISVSYDNFALV-----DYLVFERTKSDTD 25
Db 117 ISVSYDNFALVCRSIIFTSQPKEDDYLVFERTKSDTD 153
RESULT 9
AAR59718
ID AAR59718 standard; protein; 142 AA.
XX
XX AAR59718;
AC 25-MAR-2003 (revised)
DT 02-MAR-1995 (first entry)
XX
XX Thrombin inhibitory protein T15.
DE
XX Thrombin; inhibition; T112; T128; T145; T15; saliva;
XX blood sucking insect; Triatoma pallidipennis; thrombosis;
XX arteriosclerosis; unstable angina; vascular reocclusion; coagulation;
XX haemo-dialysis; serine protease.
XX
XX Triatoma pallidipennis.
OS
XX
XX WO9413807-A1.
PN
XX 23-JUN-1994.
PD
XX
XX 03-DEC-1993; 93WO-DE001172.
XX
XX 04-DEC-1992; 92DE-04241659.
PR 12-FEB-1993; 93DE-04304731.
PR 17-AUG-1993; 93DE-04328336.
PR 25-NOV-1993; 93DE-04340798.
XX
XX (SCHD ) SCHERING AG.
XX
XX Noeskejungblut C, Schleuning W, Alagon A, Possani L;
PI Cuevas-Aguirre D, Donner P, Haendler B, Hechler U;
XX

```

DR WPI; 1994-217883/26.
 DR N-PSDB; AAQ68271.
 XX
 PT New thrombin inhibitory proteins from saliva of blood sucking insects -
 PT esp. Triatoma pallidipennis, for treating or preventing thrombosis,
 PT unstable angina, vascular reocclusion etc., also corresponding DNA,
 PT vectors and transformed cells.
 XX
 PS Claim 4; Page 30; 57pp; German.
 XX
 CC Thrombin inhibitory proteins are isolated from Triatoma pallidipennis.
 CC They are used to treat thrombosis, unstable angina or arteriosclerosis;
 CC to prevent vascular reocclusion after balloon angioplasty; and to inhibit
 CC coagulation during haemo-dialysis. The DNA sequences of thrombin
 CC inhibitory proteins T112, T128, T145 and T15 are given in AAQ68272-75
 CC (sequences encoding the corresp. mature proteins are given in AAQ68268-
 CC 71). The proteins are specific for thrombin, even at a 40-fold excess
 CC other serine proteases are not inhibited. (Updated on 25-MAR-2003 to
 CC correct PN field.) (Updated on 25-MAR-2003 to correct PI field.)
 XX
 SQ Sequence 142 AA;
 Query Match 78.2%; Score 98.5; DB 2; Length 142;
 Best Local Similarity 62.2%; Pred. No. 8.1e-08;
 Matches 23; Conservative 0; Mismatches 1; Indels 13; Gaps 1;
 OY 2 ISVSYDNFALV-----DYLVFERTKSDTD 25
 Db 99 ISVSYDNFALVCRSITFTSQPKEDYLVLEERTKSDTD 135
 RESULT 10
 AARS9722
 ID AARS9722 standard; protein; 160 AA.
 XX
 AC AARS9722;
 XX
 DT 25-MAR-2003 (revised)
 DT 02-MAR-1995 (first entry)
 XX
 DE Thrombin inhibitory protein T15.
 XX
 KW Thrombin; inhibition; T112; T128; T145; T15; saliva;
 KW blood sucking insect; Triatoma pallidipennis; thrombosis;
 KW arteriosclerosis; unstable angina; vascular reocclusion; coagulation;
 KW haemo-dialysis; serine protease.
 XX
 OS Triatoma pallidipennis.
 XX
 FH Key Location/Qualifiers
 FH Peptide 1..18
 FT /label= sig_peptide
 FT 19..160
 FT Protein /label= mat_protein
 FT
 XX
 PN WO9413807-A1.
 PD 23-JUN-1994.
 XX
 PF 03-DEC-1993; 93WO-DE001172.
 XX
 PR 04-DEC-1992; 92DE-04241659.
 PR 12-FEB-1993; 93DE-04304731.
 PR 17-AUG-1993; 93DE-04328336.
 PR 25-NOV-1993; 93DE-04340798.
 XX
 PA (SCHD) SCHERING AG.
 XX
 PI Noeskejunbiut C, Schlenning W, Alagon A, Possant L;
 PI Cuevas-Aguirre D, Donner P, Haendler B, Hechler U;
 DR WPI; 1994-217883/26.
 DR N-PSDB; AAQ68275.

XX
 PT New thrombin inhibitory proteins from saliva of blood sucking insects -
 PT esp. Triatoma pallidipennis, for treating or preventing thrombosis,
 PT unstable angina, vascular reocclusion etc., also corresponding DNA,
 PT vectors and transformed cells.
 XX
 PS Claim 5; Page 34; 57pp; German.
 XX
 CC Thrombin inhibitory proteins are isolated from Triatoma pallidipennis.
 CC They are used to treat thrombosis, unstable angina or arteriosclerosis;
 CC to prevent vascular reocclusion after balloon angioplasty; and to inhibit
 CC coagulation during haemo-dialysis. The DNA sequences of thrombin
 CC inhibitory proteins T112, T128, T145 and T15 are given in AAQ68272-75
 CC (sequences encoding the corresp. mature proteins are given in AAQ68268-
 CC 71). The proteins are specific for thrombin, even at a 40-fold excess
 CC other serine proteases are not inhibited. (Updated on 25-MAR-2003 to
 CC correct PN field.) (Updated on 25-MAR-2003 to correct PI field.)
 XX
 SQ Sequence 160 AA;
 Query Match 78.2%; Score 98.5; DB 2; Length 160;
 Best Local Similarity 62.2%; Pred. No. 9.4e-08;
 Matches 23; Conservative 0; Mismatches 1; Indels 13; Gaps 1;
 OY 2 ISVSYDNFALV-----DYLVFERTKSDTD 25
 Db 117 ISVSYDNFALVCRSITFTSQPKEDYLVLEERTKSDTD 153
 RESULT 11
 ABU31592
 ID ABU31592 standard; protein; 544 AA.
 XX
 AC ABU31592;
 XX
 DT 19-JUN-2003 (first entry)
 DE Protein encoded by Prokaryotic essential gene #17119.
 XX
 KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX
 OS Klebsiella pneumoniae.
 XX
 PN WO200277183-A2.
 PD 03-OCT-2002.
 XX
 PF 21-MAR-2002; 2002WO-US009107.
 XX
 PR 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 PA (BLIT-) ELITRA PHARM INC.
 XX
 PI Wang L, Zamudio C, Malone C, Haseelbeck R, Ohlsen KL, Zykkind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX
 DR WPI; 2003-029926/02.
 DR N-PSDB; ACA35462.
 XX
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 PS Claim 25; SEQ ID NO 59516; 1766pp; English.
 XX
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid

CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-regulated gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

SQ Sequence 544 AA;

Query Match Best Local Similarity 40.5%; Score 51; DB 6; Length 544;

Matches 10; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 2 ISVSYDNFALVDYLVFERTKSDTD 25
DB 369 LNVGLDNFPLVKYLISQVMQKXD 392

RESULT 12

ABO61580 ABO61580 standard; protein; 548 AA.

AC ABO61580;

DT 29-JUL-2004 (first entry)

XX Klebsiella pneumoniae polypeptide seqid 8097.

XX Recombinant expression vector; transcription regulatory element;

KM Klebsiella pneumoniae protein; antibacterial; Vaccine.

XX Klebsiella pneumoniae.

OS US6610836-B1.

XX US6610836-B1.

PD 26-AUG-2003.

XX 27-JAN-2000; 2000US-00489039.

XX 29-JAN-1999; 99US-0117747P.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX Breton GL, Osborne M;

XX WPI; 2003-895346/82.

XX N-PSDB; ACH95131.

XX New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for
XX preparing a vaccine composition against Klebsiella pneumoniae.
XX Disclosure; SEQ ID NO 8097; 932pp; English.

XX The invention describes a new isolated nucleic acid encoding a Klebsiella
CC pneumoniae polypeptide. Also described are: a recombinant expression
CC vector comprising the nucleic acid, operably linked to a transcription
CC regulatory element; and a cell comprising the recombinant expression
CC vector. The nucleic acid is useful for preparing a vaccine composition
CC against Klebsiella pneumoniae. This is the amino acid sequence of a
CC Klebsiella pneumoniae polypeptide of the invention

SQ Sequence 548 AA;

Query Match Best Local Similarity 40.5%; Score 51; DB 7; Length 548;

Matches 10; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 2 ISVSYDNFALVDYLVFERTKSDTD 25
DB 373 LNVGLDNFPLVKYLISQVMQKXD 396

RESULT 13

ABU28453 ABU28453 standard; protein; 549 AA.

AC ABU28453;

DT 19-JUN-2003 (first entry)

XX Protein encoded by prokaryotic essential gene #13980.

XX Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX Enterobacter cloacae.

XX WO200277183-A2.

XX 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US009107.

XX 21-MAR-2001; 2001US-00815242.

XX 06-SEP-2001; 2001US-00948993.

XX 25-OCT-2001; 2001US-0342923P.

XX 08-FEB-2002; 2002US-00072851.

XX 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Foreyth RA, Xu HH;

XX WPI; 2003-029926/02.

XX N-PSDB; ACA32323.

XX New antisense nucleic acids, useful for identifying proteins or screening

XX PT for homologous nucleic acids, required for cellular proliferation to

XX PT isolate candidate molecules for rational drug discovery programs.

XX Claim 25; SEQ ID NO 56377; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)

CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 549 AA;

Query Match 39.7%; Score 50; DB 6; Length 549;
Best Local Similarity 45.8%; Pred. No. 32;
Matches 11; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

OY 2 ISVSYDNFALVDYLVFERTKSDTD 25
DB 374 MDVGLNFDLVKYLISQVMSLSDDD 397

RESULT 14

ABO67213
ID ABO67213 standard; protein; 565 AA.

XX ABO67213;

AC 29-JUL-2004 (first entry)

XX Klebsiella pneumoniae polypeptide seqid 13730.

DE Recombinant expression vector; transcription regulatory element;

KM Klebsiella pneumoniae protein; antibacterial; vaccine.

XX Klebsiella pneumoniae.

OS US6610836-B1.

XX 26-AUG-2003.

XX 27-JAN-2000; 2000US-00489039.

XX 29-JAN-1999; 99US-0117747P.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX Breton GL, Osborne M;

XX WPI; 2003-895346/82.

XX N-PSDB; ABD00784.

XX New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for
XX preparing a vaccine composition against Klebsiella pneumoniae.

XX Disclosure; SEQ ID NO 13730; 932pp; English.

XX The invention describes a new isolated nucleic acid encoding a Klebsiella
XX pneumoniae polypeptide. Also described are: a recombinant expression
XX vector comprising the nucleic acid, operably linked to a transcription
XX regulatory element; and a cell comprising the recombinant expression
XX vector. The nucleic acid is useful for preparing a vaccine composition
XX against Klebsiella pneumoniae. This is the amino acid sequence of a
XX Klebsiella pneumoniae polypeptide of the invention

SQ Sequence 565 AA;

Query Match 39.7%; Score 50; DB 7; Length 565;
Best Local Similarity 54.5%; Pred. No. 33;
Matches 12; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

OY 4 VSYDNFALVDYLVFERTKSDTD 25
DB 388 VGIDNFDLVKYLISQVMSLSDDD 409

RESULT 15

AAB36283
ID AAB36283 standard; protein; 213 AA.

XX AAB36283;

AC 21-FEB-2001 (first entry)

XX Human protein fragment PN7065 SEQ ID NO: 4.

DE Protein-protein interaction; physiological disorder; NIDDM;

KM neurodegenerative disorder; drug screening; predisposition;

XX non-insulin dependent diabetes mellitus.

XX Homo sapiens.

XX MO200065340-A1.

XX 02-NOV-2000.

XX 21-APR-2000; 2000MO-US010651.

XX 22-APR-1999; 99US-0130389P.

XX 24-JUN-1999; 99US-0140693P.

XX 30-SEP-1999; 99US-0156847P.

XX 02-NOV-1999; 99US-0163073P.

XX 02-DEC-1999; 99US-0168376P.

XX 02-DEC-1999; 99US-0168378P.

XX (MYRI-) MYRIAD GENETICS INC.

XX Heichman K, Bartel PL;

XX WPI; 2000-687376/67.

XX N-PSDB; AAC68862.

XX Novel protein complexes involved in mammalian physiological pathways, for
XX diagnosing predisposition to, or existence of diabetes or Alzheimer's
XX disease, and in drug screening for identifying modulators of complexes.

XX Claim 43; Page 47; 87pp; English.

XX The present invention relates to novel protein-protein interactions which
XX are involved in mammalian physiological diseases and disorders. These
XX include non-insulin dependent diabetes mellitus (NIDDM) and
XX neurodegenerative diseases such as Alzheimer's disease. The protein
XX complexes described can be used in the diagnosis of these diseases and a
XX predisposition to them, in drug screening and in the identification of
XX other proteins involved in the same pathway

XX Sequence 213 AA;

Query Match 38.1%; Score 48; DB 3; Length 213;
Best Local Similarity 52.9%; Pred. No. 22;
Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

OY 5 SYDNFALVDYLVFERTK 21
DB 56 SYNHFALVYTLDERLK 72

RESULT 16

ID	AA047219
XX	AA047219 standard; protein, 782 AA.
AC	XX
XX	AA047219;
DT	12-FEB-2002 (first entry)
XX	XX
DE	Human NOVI0 protein.
XX	XX
OS	Homo sapiens.
XX	XX
PB	WO200174651-A2.
PD	11-OCT-2001.
PF	30-MAR-2001; 2001WO-US010039.
PR	30-MAR-2000; 2000US-0193205P.
PR	30-MAR-2000; 2000US-0193339P.
PR	05-APR-2000; 2000US-0195343P.
PR	06-APR-2000; 2000US-0195005P.
PR	10-APR-2000; 2000US-0195088P.
PR	11-APR-2000; 2000US-0195566P.
PR	13-APR-2000; 2000US-0197081P.
PR	14-APR-2000; 2000US-0197087P.
PR	14-APR-2000; 2000US-0197525P.
PR	29-MAR-2001; 2001US-00823187.
PA	(CURA-) CURAGEN CORP.
PI	Majumder K, Spaderna SK, Taupier RJ, Padigaru M, Burgess CE;
PI	Shimkets RA, Spytek KA, Liu X, Patturajan M, Gusev VY;
DR	N-PDSB; ABA02000.
PT	New G-protein-coupled receptor related polypeptides and polymucleotides for diagnosis, prevention and treatment of metabolic, neurodegenerative, retinal, immune, hematopoietic disorders, diabetes, obesity and infections.
PS	Claim 1; Page 103; 194pp; English.
CC	The present invention provides the protein and coding sequences of novel human G-protein coupled receptors, designated NOV1, NOV2, NOV3, NOV4, NOV5a, NOV5b, NOV6a, NOV6b, NOV7, NOV8a, NOV8b, NOV9 and NOV10. These can be used in the treatment of NOVX related diseases, including cancer, metabolic, neurodegenerative, immune, haematopoietic, developmental, CC retinal, feeding, neurological and psychotic diseases and disorders and CC infections. The present sequence is the NOV10 protein, which shares homology with the salt-inducible protein kinase
SQ	Sequence 782 AA:
Query Match	38.1%; Score 48; DB 4; Length 782;
Best Local Similarity	52.9%; Pred. No. 1e+02;
Matches	9; Conservative 4; Mismatches 4; Indels 0; Gaps 0,

Db		330 SYNHPAIIYLLERLK 346
RESULT 17		
ID	AAG65764	standard; protein; 783 AA.
XX		
AC	AAG65764;	
XX		
DT	07-JAN-2002	(first entry)
XX		
DE	Human protein kinase polypeptide 3714.	
XX		
KM	Protein kinase; 3714; 16742; 23546; 13887; cancer; bone disorder; human; cytosolic; antiinflammatory; immunosuppressive; cardiac; hepatotropic; gene therapy.	
XX		
OS	Homo sapiens.	
XX		
PN	WO200173050-A2.	
PD	04-OCT-2001.	
XX		
PF	23-MAR-2001; 2001WO-US009483.	
XX		
PR	24-MAR-2000; 2000US-0191846P.	
XX		
PA	(MILL-) MILLENNIUM PHARM INC.	
PI	Meyers R;	
XX		
DR	WPI; 2001-611632/70.	
XX		
DR	N-PSDB; AAI66822; AAI66823.	
XX		
PT	New human protein kinase polypeptides, 3714, 16742, 23546 and 13887,	
PT	useful in diagnosis of cancer or cellular proliferation or	
PT	differentiation disorders and to screen for polypeptide modulators useful	
PT	to treat such conditions.	
XX		
PS	Claim 4; Fig 1A-D; 169pp; English.	
XX		
CC	The invention provides novel human protein kinase polypeptides, 3714,	
CC	16742, 23546 and 13887 and nucleic acid molecules encoding them. The	
CC	protein kinase polypeptides can be expressed by standard recombinant	
CC	methodology. 3714, 16742, 23546 or 13887 nucleic acids and polypeptides	
CC	are useful for diagnostic and screening methods to identify subjects (at	
CC	risk of) having cancer or cellular proliferation and/or differentiation	
CC	disorders. 3714, 16742, 23546 or 13887 nucleic acids, polypeptides and	
CC	modulators are useful for the treatment of cancer, particularly colon	
CC	cancer or cellular proliferation and/or differentiation disorders. Other	
CC	disorders associated with 3714, 16742, 23546 or 13887 expression or	
CC	activity that can be treated include bone related disorders, inflammatory	
CC	disorders, autoimmune diseases, cardiovascular disorders and liver	
CC	diseases. The present sequence represents a human protein kinase	
CC	polypeptide 3714	
XX		
SQ	Sequence 783 AA;	
Query Match	38.1%; Score 48; DB 4; Length 783;	
Best Local Similarity	52.9%; Pred. No. 1e+02;	
Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0		
OY	5 SYDNFALVDLVLPFRTK 21 :: : : 330 SYNHPAIIYLLERLK 346	
Db		
RESULT 18		
ID	AAB85786	
XX		
XX	AAB85786 standard; protein; 783 AA.	
XX		

PF 06-SEP-1999; 99WO-DE002865.
XX
PR 19-SEP-1998; 98DE-01042863.
XX
PA (KRAM/) KRAMER M.
XX
PI Kramer M, Bechtel M, Reinartz J, Schaefer B, Walllich R;
XX WPI: 2000-283542/24.
DR N-PSDB; AAX82951.
XX
PT New regulatory polypeptide from keratinocytes, useful for diagnosis and
PT treatment of dermatological disorders, also in cosmetics.
XX
PS Claim 1; Page 51; 55pp; German.
XX
CC This invention describes a novel human regulatory polypeptide designated
CC pK#122 (I), or its variants from human keratinocytes, which in activated
CC keratinocytes is the same as, or similar to, the most highly expressed
CC protein and has antiproliferative activity. (I) is a kinase involved in
CC signal transduction and is conjectured to be involved in cell-cell or
CC cell-matrix interactions and/or formation of (hem)desmosomes. Modulation
CC of (I) will affect proliferation and differentiation of keratinocytes.
CC (I) is upregulated in some dermatological conditions, e.g. pemphigus
CC vulgaris and psoriasis vulgaris. (I) is used to raise specific antibodies
CC (Ab) which are used for diagnosis and treatment of dermatological
CC disorders and for cosmetic treatment of the epidermis. Nucleic acid (II)
CC that encodes (I) is used; (I) to produce transgenic mammals (particularly
CC mice and rats); and (II) as source of sense and antisense
CC oligonucleotides for diagnosis and treatment of dermatological disorders
CC and for cosmetic treatment of the epidermis. (II), (I) and polypeptides
CC (Ia) produced by splice variants of (II) are used to identify specific
CC binding agents (potentially useful in medicine and cosmetics) that
CC modulate their function and/or expression, especially inhibitors and
CC activators. Cells and animals transformed with (II) are used to study (I)
CC -induced changes in cell morphology and basic functions. This sequence
CC represents the human keratinocyte protein pK#122 described in the method
CC of the invention
XX
SQ Sequence 790 AA;
XX
Query Match 38.1%; Score 48; DB 3; Length 790;
Best Local Similarity 52.9%; Pred. No. 1e+02;
Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
XX
QY 5 SYDNFALVDYVFERK 21
DB 337 SYNHPALYYLLERLK 353
XX
RESULT 21
ABM84424
ID ABM84424 standard; protein; 818 AA.
XX
AC ABM84424;
XX
DT 18-NOV-2004 (first entry)
XX
DE Human diagnostic and therapeutic pprotein SEQ ID NO:4673.
XX
KW gene therapy; human diagnostic and therapeutic polynucleotide; dithp.
XX
OS Homo sapiens.
XX
PN WO2004023973-A2.
XX
PD 25-MAR-2004.
XX
PF 12-SEP-2003; 2003WO-US028227.
XX
PR 12-SEP-2002; 2002US-0410259P.
XX
PR 12-SEP-2002; 2002US-0410260P.
XX

PA (INCY-) INCYTE CORP.
XX
PI Schmidt JP, Wright RJ, Bruns CM, Marjanovic MW, Shen F;
PI Hartshorne TA, Suchanowski MT, Altus CM, Pitts SJ, Elder LV;
PI Mooney EM, Deleage AM, Panesar IS, Banville SC, Reddy TP;
PI Stevens KA, Blanchard JM, Panzer SR, Wang X, Au AP, Gerstin EH;
PI Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LJ;
PI Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vlett UA, Kirtan ES;
PI Xu Y, Kwong M, Pollocky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;
PI Patury S, Shi X, Suarez CJ;
XX
DR WPI: 2004-329368/30.
DR N-PSDB; ACN43076.
XX
PT New diagnostic and therapeutic polynucleotides and polypeptides, useful
PT in diagnosing a condition, disease or disorder associated with human
PT molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or
PT in gene mapping.
XX
PS Claim 27; Page; 190pp; English.
XX
CC The invention relates to novel diagnostic and therapeutic polynucleotides
CC selected from one of the 2722 sequences defined in the specification. A
CC polynucleotide of the invention may have a use in gene therapy. The human
CC diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be
CC used to diagnose a particular condition, disease or disorder associated
CC with human molecules, e.g. cell proliferative disorders,
CC autoimmune/inflammatory disorder, developmental disorder, endocrine
CC disorder, neurological disorders, gastrointestinal disorders, or
CC infections caused by virus, bacteria, fungi or parasite. The dithp
CC molecules may also be used in genetic mapping, in identifying individuals
CC from minute biological samples, in detecting single nucleotide
CC polymorphisms, as molecular weight markers, and for somatic or germline
CC gene therapy. The present sequence represents a dithp protein of the
CC invention. Note: The sequence data for this patent is not represented in
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at www.wipo.int/pct/en/sequences/listing.htm
XX
SQ Sequence 818 AA;
XX
Query Match 38.1%; Score 48; DB 8; Length 818;
Best Local Similarity 52.9%; Pred. No. 1.1e+02;
Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
XX
QY 5 SYDNFALVDYVFERK 21
DB 366 SYNHPALYYLLERLK 382
XX
RESULT 22
AAW90879
ID AAW90879 standard; protein; 823 AA.
XX
AC AAW90879;
XX
DT 07-JUL-2000 (first entry)
XX
DE Human keratinocyte derived pK#122 protein #2.
XX
KW Keratinocyte; regulatory protein; human; pK#122; antiproliferative;
KW kinase; signal transduction; desmosome; dermatological condition;
KW pemphigus vulgaris; psoriasis vulgaris; epidermis; cosmetic.
XX
OS Homo sapiens.
XX
PN WO200017232-A2.
XX
PD 30-MAR-2000.
XX
PF 06-SEP-1999; 99WO-DE002865.
XX
PR 19-SEP-1998; 98DE-01042863.
XX

PA (KRAM/) KRAMER M.
XX
PI Kramer M., Bechtel M., Reinartz J., Schaefer B., Wallich R.
XX
XX WPI; 2000-283542/24.
DR N-PSDB; AAX82952.
XX
PT New regulatory polypeptide from keratinocytes, useful for diagnosis and
PT treatment of dermatological disorders, also in cosmetics.
XX
PS Claim 1; Page 53; 55pp; German.
XX
CC This invention describes a novel human regulatory polypeptide designated
CC pKeh122 (I), or its variants from human keratinocytes, which in activated
CC keratinocytes is the same as, or similar to, the most highly expressed
CC protein and has antiproliferative activity. (I) is a kinase involved in
CC signal transduction and is conjectured to be involved in cell-cell or
CC cell-matrix interactions and/or formation of (hem)desmosomes. Modulation
CC of (I) will affect proliferation and differentiation of keratinocytes.
CC (I) is upregulated in some dermatological conditions, e.g. pemphigus
CC vulgaris and psoriasis vulgaris. (I) is used to raise specific antibodies
CC (Ab) which are used for diagnosis and treatment of dermatological
CC disorders and for cosmetic treatment of the epidermis. Nucleic acid (II)
CC that encodes (I) is used; (I) to produce transgenic mammals (particularly
CC mice and rats); and (II) as source of sense and antisense
CC oligonucleotides for diagnosis and treatment of dermatological disorders
CC and for cosmetic treatment of the epidermis. (II), (I) and polypeptides
CC (Ia) produced by splice variants of (II) are used to identify specific
CC binding agents (potentially useful in medicine and cosmetics) that
CC modulate their function and/or expression, especially inhibitors and
CC activators. Cells and animals transformed with (II) are used to study (I)
CC -induced changes in cell morphology and basic functions. This sequence
CC represents the human keratinocyte protein pKeh122 described in the method
CC of the invention
XX
SQ Sequence 823 AA;
XX
Query Match 38.1%; Score 48; DB 3; Length 823;
Best Local Similarity 52.9%; Pred. No. 1.1e+02;
Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
XX
QY 5 SYDNFALVDYLVFERK 21
DB 370 SYNHFALYYLLERLK 386
XX
RESULT 23
ABP96069 standard; protein; 1078 AA.
XX
AC ABP96069;
XX
DT 07-MAY-2003 (first entry)
XX
DE Human protein kinase SEQ ID NO:44.
XX
KM Human; protein kinase; enzyme; antiasthmatic; antiinflammatory;
KM antidiabetic; antiparkinsonian; antimigraine; cardiant; cyostatic;
KM immunosuppressive; vulnery; gene therapy; COPD; asthma; migraine;
KM chronic obstructive pulmonary disease; non-insulin dependent diabetes;
KM Parkinson's disease; myocardial infarction; inflammatory bowel disease;
KM autoimmune disorder; allograft rejection; graft versus host disease;
KM cancer; leukaemia; wound granulation.
XX
OS Homo sapiens.
XX
PN WO2003000901-A2.
XX
PD 03-JAN-2003.
XX
PF 24-JUN-2002; 2002WO-IB002358.
XX
PT 26-JUN-2001; 2001US-0301098P.

PR 06-NOV-2001; 2001US-0332870P.
XX
PA (DECO-) DECODE GENETICS EHF.
XX
PI Martinez RM, Sigurdson GT.
XX
XX WPI; 2003-201429/19.
DR N-PSDB; AB277147.
XX
PT New protein kinase genes and polypeptides, useful for diagnosing diseases
PT associated with a protein kinase, or in gene therapy for treating e.g.
PT Parkinson's disease, migraine, myocardial infarction, allograft rejection
PT or cancers.
XX
PS Claim 9; Page 71-72; 258pp; English.
XX
CC AB277126 to AB277165 encode the human protein kinases given in ABP96048
CC to ABP96087. The protein kinases have antiasthmatic, antiinflammatory,
CC antidiabetic, antiparkinsonian, antimigraine, cardiant, cyostatic,
CC immunosuppressive and vulnery activities, and can be used in gene
CC therapy. A protein kinase therapeutic agent from the present invention,
CC particularly a protein kinase gene agonist or antagonist, can be used for
CC treating a disease or condition associated with a protein kinase in an
CC individual. These diseases include chronic obstructive pulmonary diseases
CC (COPD), asthma, non-insulin dependent diabetes, Parkinson's disease,
CC migraine, myocardial infarction, inflammatory bowel disease, autoimmune
CC disorders (e.g. allograft rejection or graft vs. host disease), cancers
CC (e.g. leukaemia) or wound granulation
XX
SQ Sequence 1078 AA;
XX
Query Match 38.1%; Score 48; DB 6; Length 1078;
Best Local Similarity 52.9%; Pred. No. 1.5e+02;
Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
XX
QY 5 SYDNFALVDYLVFERK 21
DB 395 SYNHFALYYLLERLK 411
XX
RESULT 24
ADC95348 standard; protein; 321 AA.
XX
AC ADC95348;
XX
DT 01-JAN-2004 (first entry)
XX
DE E. faecium protein sequence SEQ ID 4975.
XX
KM Vaccine; urinary tract infection; bacteraemia; endocarditis; wound;
KM abdominal-pelvic infection.
XX
OS Enterococcus faecium.
XX
PN US6583275-B1.
XX
PD 24-JUN-2003.
XX
PF 30-JUN-1998; 98US-00107532.
XX
PR 02-JUL-1997; 97US-0051571P.
PR 14-MAY-1998; 98US-0085598P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Doucette-Stamm LA, Bush D;
XX
XX WPI; 2003-799836/75.
DR N-PSDB; ADC91694.
XX
PT New isolated nucleic acid derived from Enterococcus faecium encoding an
PT Enterococcus faecium polypeptide useful for detection, prevention and

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2005, 12:19:33 ; Search time 42 Seconds
(without alignments)
44.434 Million cell updates/sec

Title: US-10-694-847-3
Perfect score: 126
Sequence: 1 KISVSYDNFALVDYLVFERTKSDPTD 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues
Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/1/aa/5A_COMB.pep:*
2: /cgn2_6/prodata/1/aa/5B_COMB.pep:*
3: /cgn2_5/prodata/1/aa/6A_COMB.pep:*
4: /cgn2_6/prodata/1/aa/6B_COMB.pep:*
5: /cgn2_6/prodata/1/aa/PCTUS_COMB.pep:*
6: /cgn2_6/prodata/1/aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	104.5	82.9	142	2	US-08-448-438-1	Sequence 1, Appl1
2	104.5	82.9	142	2	US-08-448-438-2	Sequence 2, Appl1
3	104.5	82.9	142	2	US-08-448-438-3	Sequence 3, Appl1
4	104.5	82.9	160	2	US-08-448-438-5	Sequence 5, Appl1
5	104.5	82.9	160	2	US-08-448-438-6	Sequence 6, Appl1
6	104.5	82.9	160	2	US-08-448-438-7	Sequence 7, Appl1
7	98.5	78.2	142	2	US-08-448-438-4	Sequence 4, Appl1
8	98.5	78.2	160	2	US-08-448-438-8	Sequence 8, Appl1
9	51	40.5	548	4	US-09-489-039A-8097	Sequence 8097, Ap
10	50	39.7	555	4	US-09-489-039A-13730	Sequence 13730, A
11	47.5	37.7	321	4	US-09-107-532A-4975	Sequence 4975, Ap
12	47	37.3	330	4	US-09-538-092-180	Sequence 180, App
13	46.5	36.9	760	4	US-08-195-152-2	Sequence 2, Appl1
14	46.5	36.9	760	4	US-08-754-311B-2	Sequence 2, Appl1
15	46	36.5	529	4	US-09-801-042-2	Sequence 2, Appl1
16	45	35.7	477	4	US-09-523-849-34	Sequence 34, Appl
17	44.5	35.3	477	4	US-09-134-000C-5553	Sequence 5553, A
18	44	34.9	154	4	US-09-270-767-40476	Sequence 40476, A
19	44	34.9	154	4	US-09-270-767-55692	Sequence 55692, A
20	44	34.9	293	4	US-09-248-796A-15099	Sequence 15099, A
21	43	34.1	89	4	US-09-134-000C-3521	Sequence 3521, Ap
22	43	34.1	232	3	US-09-066-408-6	Sequence 6, Appl1
23	43	34.1	235	4	US-09-270-767-47930	Sequence 47930, A
24	43	34.1	283	4	US-09-107-532A-5418	Sequence 5418, Ap
25	43	34.1	430	1	US-08-601-435-2	Sequence 2, Appl1
26	43	34.1	430	2	US-08-931-047-2	Sequence 2, Appl1
27	43	34.1	430	2	US-08-783-202-2	Sequence 2, Appl1

28	43	34.1	430	4	US-09-443-041A-31	Sequence 31, Appl
29	43	34.1	527	4	US-09-107-532A-5945	Sequence 5945, Ap
30	42.5	33.7	131	4	US-09-621-976-5247	Sequence 5247, Ap
31	42.5	33.7	146	4	US-09-621-976-5248	Sequence 5248, Ap
32	42.5	33.7	169	4	US-09-770-834-7	Sequence 7, Appl1
33	42	33.3	31	4	US-09-155-203-3	Sequence 3, Appl1
34	42	33.3	38	4	US-09-155-203-2	Sequence 2, Appl1
35	42	33.3	39	4	US-09-155-203-1	Sequence 1, Appl1
36	42	33.3	84	4	US-09-134-000C-6573	Sequence 6573, Ap
37	42	33.3	215	3	US-09-305-984-64	Sequence 64, Appl
38	42	33.3	215	3	US-09-493-940-64	Sequence 64, Appl
39	42	33.3	224	3	US-09-305-984-72	Sequence 72, Appl
40	42	33.3	224	3	US-09-305-984-74	Sequence 74, Appl
41	42	33.3	224	3	US-09-305-984-76	Sequence 76, Appl
42	42	33.3	224	4	US-09-493-940-74	Sequence 74, Appl
43	42	33.3	224	4	US-09-493-940-76	Sequence 76, Appl
44	42	33.3	224	4	US-09-493-940-78	Sequence 78, Appl
45	42	33.3	453	4	US-09-248-796A-14922	Sequence 14922, A
46	42	33.3	524	4	US-09-489-039A-12626	Sequence 12626, A
47	42	33.3	550	4	US-09-328-352-5727	Sequence 5727, Ap
48	42	33.3	664	4	US-09-621-816B-10	Sequence 10, Appl
49	42	33.3	665	4	US-09-621-816B-6	Sequence 6, Appl1
50	42	33.3	887	1	US-07-867-106-3	Sequence 3, Appl1
51	42	33.3	1170	1	US-08-313-288B-20	Sequence 20, Appl1
52	42	33.3	1170	4	US-09-657-472-2	Sequence 2, Appl1
53	41.5	32.9	63	4	US-09-248-796A-27007	Sequence 27007, A
54	41.5	32.9	215	2	US-08-836-443-4	Sequence 4, Appl1
55	41.5	32.9	445	4	US-09-252-991A-20277	Sequence 20277, A
56	41.5	32.9	751	2	US-08-836-443-3	Sequence 3, Appl1
57	41.5	32.9	953	4	US-09-949-016-8262	Sequence 8262, Ap
58	41.5	32.9	968	4	US-09-949-016-8262	Sequence 8262, Ap
59	41	32.5	61	4	US-09-248-796A-21384	Sequence 21384, A
60	41	32.5	109	4	US-09-513-999-4500	Sequence 4500, Ap
61	41	32.5	115	4	US-09-270-767-33744	Sequence 33744, A
62	41	32.5	115	4	US-09-270-767-33744	Sequence 33744, A
63	41	32.5	129	3	US-09-342-647-22	Sequence 48961, A
64	41	32.5	215	3	US-09-305-984-18	Sequence 22, Appl
65	41	32.5	215	4	US-09-073-541A-18	Sequence 18, Appl
66	41	32.5	215	4	US-09-493-940-18	Sequence 18, Appl
67	41	32.5	215	4	US-09-583-110-5320	Sequence 5320, Ap
68	41	32.5	219	4	US-09-107-433-4810	Sequence 4810, Ap
69	41	32.5	221	4	US-09-107-532A-6667	Sequence 6667, Ap
70	41	32.5	221	4	US-09-134-000C-6517	Sequence 6517, Ap
71	41	32.5	240	4	US-09-489-039A-8615	Sequence 8615, Ap
72	41	32.5	281	4	US-09-248-796A-23359	Sequence 23359, A
73	41	32.5	455	2	US-08-870-827-3	Sequence 3, Appl1
74	41	32.5	455	3	US-09-317-179-3	Sequence 3, Appl1
75	41	32.5	455	4	US-09-986-682B-3	Sequence 3, Appl1
76	41	32.5	457	4	US-09-134-000C-6461	Sequence 6461, Ap
77	41	32.5	635	2	US-09-014-969-11	Sequence 11, Appl
78	41	32.5	635	4	US-09-949-016-6991	Sequence 6991, Ap
79	41	32.5	666	4	US-09-949-016-7874	Sequence 7874, Ap
80	40	31.7	149	4	US-09-198-452A-792	Sequence 792, App
81	40	31.7	149	4	US-09-438-185A-744	Sequence 744, App
82	40	31.7	217	4	US-09-107-532A-3702	Sequence 3702, Ap
83	40	31.7	234	4	US-09-107-532A-6523	Sequence 6523, Ap
84	40	31.7	234	4	US-09-107-532A-5072	Sequence 5072, Ap
85	40	31.7	271	4	US-09-107-532A-5071	Sequence 5071, Ap
86	40	31.7	275	4	US-09-248-796A-25216	Sequence 25216, A
87	40	31.7	347	4	US-09-540-236-3086	Sequence 3086, Ap
88	40	31.7	360	4	US-09-583-110-3285	Sequence 3285, Ap
89	40	31.7	361	4	US-09-107-433-3390	Sequence 3390, Ap
90	40	31.7	446	4	US-09-591-466C-4	Sequence 4, Appl1
91	40	31.7	514	4	US-09-711-164-326	Sequence 326, App
92	40	31.7	518	4	US-09-489-039A-8407	Sequence 8407, Ap
93	40	31.7	548	4	US-09-134-000C-4827	Sequence 4827, Ap
94	40	31.7	644	4	US-09-248-796A-16454	Sequence 16454, A
95	40	31.7	734	4	US-09-134-000C-55602	Sequence 55602, Ap
96	40	31.7	755	4	US-09-975-326-4	Sequence 4, Appl1
97	40	31.7	765	4	US-10-217-357-4	Sequence 4, Appl1
98	40	31.7	766	4	US-09-975-326-2	Sequence 2, Appl1
99	40	31.7	766	4	US-10-217-357-2	Sequence 2, Appl1
100	40	31.7	789	4	US-09-252-991A-27011	Sequence 27011, A

ALIGNMENTS

RESULT 1
US-08-448-438-1; Sequence 1, Application US/08448438
; Patent No. 5876971

GENERAL INFORMATION:

APPLICANT: NOESKE-JUNGBLUT, CHRISTIANE

APPLICANT: SCHLEUNING, WOLF-DIETER

APPLICANT: ALAGON, ALEJANDRO

APPLICANT: POSSANI, LOURIVAL

APPLICANT: CUEVAS-AGUIRRE, DELIA

APPLICANT: DONNER, PETER

APPLICANT: HAENDLER, BERNARD

TITLE OF INVENTION: THROMBIN INHIBITOR FROM THE SALIVA OF

NUMBER OF SEQUENCES: 29

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/448,438

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/460,383

FILING DATE: 02-JUN-1995

APPLICATION NUMBER: WO PCT/DE93/91172

FILING DATE: 03-DEC-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: DE 4340798.6

FILING DATE: 25-NOV-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: DE 4328336.5

FILING DATE: 17-AUG-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: DE 4304731.9

FILING DATE: 12-FEB-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: DE 4241659.0

FILING DATE: 04-DEC-1992

ATTORNEY/AGENT INFORMATION:

NAME: Hamlet-King, Diana

REGISTRATION NUMBER: 33,302

REFERENCE/DOCKET NUMBER: SCH 1474C1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 703-243-6333

TELEFAX: 703-243-6410

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 142 amino acids

TYPE: amino acid sequence

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

ORGANISM: SALIVA GLAND OF TRIATOMA PALIDIPENNIS

INDIVIDUAL ISOLATE: T1 12

US-08-448-438-1

Query Match 82.9%; Score 104.5; DB 2; Length 142;
Best Local Similarity 64.9%; Pred. No. 2,4e-09;
Matches 24; Conservative 0; Mismatches 0; Indels 13; Gaps 1;QY 2 ISVSYNFALV-----DYLVEFRTKSDTD 25
DB 99 ISVSYNFALVCRSIIFTSQPKEDYLVFERTKSDTD 135RESULT 2
US-08-448-438-2; Sequence 2, Application US/08448438
; Patent No. 5876971

GENERAL INFORMATION:

APPLICANT: NOESKE-JUNGBLUT, CHRISTIANE

APPLICANT: SCHLEUNING, WOLF-DIETER

APPLICANT: ALAGON, ALEJANDRO

APPLICANT: POSSANI, LOURIVAL

APPLICANT: CUEVAS-AGUIRRE, DELIA

APPLICANT: DONNER, PETER

APPLICANT: HAENDLER, BERNARD

TITLE OF INVENTION: THROMBIN INHIBITOR FROM THE SALIVA OF

NUMBER OF SEQUENCES: 29

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/448,438

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/460,383

FILING DATE: 02-JUN-1995

APPLICATION NUMBER: WO PCT/DE93/91172

FILING DATE: 03-DEC-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: DE 4340798.6

FILING DATE: 25-NOV-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: DE 4328336.5

FILING DATE: 17-AUG-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: DE 4304731.9

FILING DATE: 12-FEB-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: DE 4241659.0

FILING DATE: 04-DEC-1992

ATTORNEY/AGENT INFORMATION:

NAME: Hamlet-King, Diana

REGISTRATION NUMBER: 33,302

REFERENCE/DOCKET NUMBER: SCH 1474C1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 703-243-6333

TELEFAX: 703-243-6410

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 142 amino acids

TYPE: amino acid sequence

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

ORGANISM: SALIVA GLAND OF TRIATOMA PALIDIPENNIS

INDIVIDUAL ISOLATE: T1 28

US-08-448-438-2

Query Match 82.9%; Score 104.5; DB 2; Length 142;
Best Local Similarity 64.9%; Pred. No. 2,4e-09;
Matches 24; Conservative 0; Mismatches 0; Indels 13; Gaps 1;QY 2 ISVSYNFALV-----DYLVEFRTKSDTD 25
DB 99 ISVSYNFALVCRSIIFTSQPKEDYLVFERTKSDTD 135

RESULT 3

US-08-448-438-3
; Sequence 3, Application US/08448438
; Patent No. 5876971
; GENERAL INFORMATION:
; APPLICANT: NOESKE-JUNGBLUT, CHRISTIANE
; APPLICANT: SCHLEUNING, WOLF-DIETER
; APPLICANT: ALAGON, ALEJANDRO
; APPLICANT: POSSANI, LOURIVAL
; APPLICANT: CUEVAS-AGUIRRE, DELIA
; APPLICANT: DONNER, PETER
; APPLICANT: HAENDLER, BERNARD
; APPLICANT: HECHLER, ULRICH
; TITLE OF INVENTION: THROMBIN INHIBITOR FROM THE SALIVA OF
; TITLE OF INVENTION: PROTOSTOMIA
; NUMBER OF SEQUENCES: 29
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/448,438
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/460,383
; FILING DATE: 02-JUN-1995
; APPLICATION NUMBER: WO PCT/DE93/91172
; FILING DATE: 03-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 4340798.6
; FILING DATE: 25-NOV-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 4328336.5
; FILING DATE: 17-AUG-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 4304731.9
; FILING DATE: 12-FEB-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 4241659.0
; FILING DATE: 04-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Hamlet-King, Diana
; REGISTRATION NUMBER: 33,302
; REFERENCE/DOCKET NUMBER: SCH 1474C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-243-6333
; TELEFAX: 703-243-6410
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 142 amino acids
; TYPE: amino acid sequence
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: SALIVA GLAND OF TRIATOMA PALPIDIPENNIS
; INDIVIDUAL ISOLATE: T1 45
; US-08-448-438-3

Query Match 82.9%; Score 104.5; DB 2; Length 142;
Best Local Similarity 64.9%; Pred. No. 2.4e-09;
Matches 24; Conservative 0; Mismatches 0; Indels 13; Gaps 1;

QY 2 ISVSYDNFALV-----DYLVEPRTKSDTD 25
Db 99 ISVSYDNFALVCRSIFTSQPKEDDYLVEPRTKSDTD 135

RESULT 4
US-08-448-438-5
; Sequence 5, Application US/08448438
; Patent No. 5876971

GENERAL INFORMATION:
; APPLICANT: NOESKE-JUNGBLUT, CHRISTIANE
; APPLICANT: SCHLEUNING, WOLF-DIETER
; APPLICANT: ALAGON, ALEJANDRO
; APPLICANT: POSSANI, LOURIVAL
; APPLICANT: CUEVAS-AGUIRRE, DELIA
; APPLICANT: DONNER, PETER
; APPLICANT: HAENDLER, BERNARD
; APPLICANT: HECHLER, ULRICH
; TITLE OF INVENTION: THROMBIN INHIBITOR FROM THE SALIVA OF
; TITLE OF INVENTION: PROTOSTOMIA
; NUMBER OF SEQUENCES: 29
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/448,438
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/460,383
; FILING DATE: 02-JUN-1995
; APPLICATION NUMBER: WO PCT/DE93/91172
; FILING DATE: 03-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 4340798.6
; FILING DATE: 25-NOV-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 4328336.5
; FILING DATE: 17-AUG-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 4304731.9
; FILING DATE: 12-FEB-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 4241659.0
; FILING DATE: 04-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Hamlet-King, Diana
; REGISTRATION NUMBER: 33,302
; REFERENCE/DOCKET NUMBER: SCH 1474C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-243-6333
; TELEFAX: 703-243-6410
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 160 amino acids
; TYPE: amino acid sequence
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: SALIVA GLAND OF TRIATOMA PALPIDIPENNIS
; INDIVIDUAL ISOLATE: T1 12
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 19..160
; US-08-448-438-5

Query Match 82.9%; Score 104.5; DB 2; Length 160;
Best Local Similarity 64.9%; Pred. No. 2.8e-09;
Matches 24; Conservative 0; Mismatches 0; Indels 13; Gaps 1;

QY 2 ISVSYDNFALV-----DYLVEPRTKSDTD 25
Db 117 ISVSYDNFALVCRSIFTSQPKEDDYLVEPRTKSDTD 153

RESULT 5
US-08-448-438-6
; Sequence 6, Application US/08448438
; Patent No. 5876971

GENERAL INFORMATION:
APPLICANT: NOESKE-JUNGBLUT, CHRISTIANE
APPLICANT: SCHLEUNING, WOLF-DIETER
APPLICANT: ALAGON, ALEJANDRO
APPLICANT: POSSANI, LOURIVAL
APPLICANT: CUEVAS-AGUIRRE, DELIA
APPLICANT: DONNER, PETER
APPLICANT: HAENDLER, BERNARD
APPLICANT: HECHLER, ULRICH
TITLE OF INVENTION: THROMBIN INHIBITOR FROM THE SALIVA OF
TITLE OF INVENTION: PROTOSTOMIA
NUMBER OF SEQUENCES: 29
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/448,438
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/460,383
FILING DATE: 02-JUN-1995
APPLICATION NUMBER: WO PCT/DE93/91172
FILING DATE: 03-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE 4340798.6
FILING DATE: 25-NOV-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE 4328336.5
FILING DATE: 17-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE 4304731.9
FILING DATE: 12-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE 4241659.0
FILING DATE: 04-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: Hamlet-King, Diana
REGISTRATION NUMBER: 33,302
REFERENCE/DOCKET NUMBER: SCH 1474C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-243-6333
TELEFAX: 703-243-6410
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 160 amino acids
TYPE: amino acid sequence
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: SALIVA GLAND OF TRIATOMA PALIDIPENNIS
INDIVIDUAL ISOLATE: T1 28
FEATURE:
NAME/KEY: Protein
LOCATION: 19..160
US-08-448-438-6

Query Match 82.9%; Score 104.5; DB 2; Length 160;
Best Local Similarity 64.9%; Pred. No. 2.8e-09;
Matches 24; Conservative 0; Mismatches 0; Indels 13; Gaps 1;

QY 2 ISVSYNFALV-----DYLVERTKSDTD 25
Db 117 ISVSYNFALVCRSITFTSQPKEDYLVFERTKSDTD 153

RESULT 6
US-08-448-438-7
Sequence 7, Application US/08448438
Patent No. 5876971

GENERAL INFORMATION:
APPLICANT: NOESKE-JUNGBLUT, CHRISTIANE
APPLICANT: SCHLEUNING, WOLF-DIETER
APPLICANT: ALAGON, ALEJANDRO
APPLICANT: POSSANI, LOURIVAL
APPLICANT: CUEVAS-AGUIRRE, DELIA
APPLICANT: DONNER, PETER
APPLICANT: HAENDLER, BERNARD
APPLICANT: HECHLER, ULRICH
TITLE OF INVENTION: THROMBIN INHIBITOR FROM THE SALIVA OF
TITLE OF INVENTION: PROTOSTOMIA
NUMBER OF SEQUENCES: 29
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/448,438
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/460,383
FILING DATE: 02-JUN-1995
APPLICATION NUMBER: WO PCT/DE93/91172
FILING DATE: 03-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE 4340798.6
FILING DATE: 25-NOV-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE 4328336.5
FILING DATE: 17-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE 4304731.9
FILING DATE: 12-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE 4241659.0
FILING DATE: 04-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: Hamlet-King, Diana
REGISTRATION NUMBER: 33,302
REFERENCE/DOCKET NUMBER: SCH 1474C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-243-6333
TELEFAX: 703-243-6410
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 160 amino acids
TYPE: amino acid sequence
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: SALIVA GLAND OF TRIATOMA PALIDIPENNIS
INDIVIDUAL ISOLATE: T1 45
FEATURE:
NAME/KEY: Protein
LOCATION: 19..160
US-08-448-438-7

Query Match 82.9%; Score 104.5; DB 2; Length 160;
Best Local Similarity 64.9%; Pred. No. 2.8e-09;
Matches 24; Conservative 0; Mismatches 0; Indels 13; Gaps 1;

QY 2 ISVSYNFALV-----DYLVERTKSDTD 25
Db 117 ISVSYNFALVCRSITFTSQPKEDYLVFERTKSDTD 153

RESULT 7
US-08-448-438-4
Sequence 4, Application US/08448438
Patent No. 5876971

GENERAL INFORMATION:
APPLICANT: NOESKE-JUNGBLUT, CHRISTIANE
APPLICANT: SCHLEUNING, WOLF-DIETER
APPLICANT: ALAGON, ALEJANDRO
APPLICANT: POSSANI, LOURIVAL
APPLICANT: CUEVAS-AGUIRRE, DELIA
APPLICANT: HECHLER, ULRICH
APPLICANT: HAENDLER, BERNARD
APPLICANT: HECHLER, ULRICH
TITLE OF INVENTION: THROMBIN INHIBITOR FROM THE SALIVA OF
TITLE OF INVENTION: PROSTOMIA
NUMBER OF SEQUENCES: 29
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/448,438
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/460,383
FILING DATE: 02-JUN-1995
APPLICATION NUMBER: WO PCT/DE93/91172
FILING DATE: 03-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE 4340798.6
FILING DATE: 25-NOV-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE 4328336.5
FILING DATE: 17-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE 4304731.9
FILING DATE: 12-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE 4241659.0
FILING DATE: 04-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: Hamlet-King, Diana
REGISTRATION/DOCKET NUMBER: 33,302
REFERENCE/DOCKET NUMBER: SCH 1474C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-243-6333
TELEFAX: 703-243-6410
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 142 amino acids
TYPE: amino acid sequence
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: SALIVA GLAND OF TRIATOMA PALIDIPENNIS
INDIVIDUAL ISOLATE: T1 5
US-08-448-438-4

Query Match 78.2%; Score 98.5; DB 2; Length 142;
Best Local Similarity 62.2%; Pred. No. 2.3e-08;
Matches 23; Conservative 0; Mismatches 1; Indels 13; Gaps 1;

QY 2 ISVSYNFALV-----DYVPERKSDTD 25
Db 99 ISVSYNFALVCRSITFTSQPKEDYLVLERTKSDTD 135

RESULT 8
US-08-448-438-8
Sequence 8, Application US/08448438
Patent No. 5876971
GENERAL INFORMATION:
APPLICANT: NOESKE-JUNGBLUT, CHRISTIANE
APPLICANT: SCHLEUNING, WOLF-DIETER

APPLICANT: ALAGON, ALEJANDRO
APPLICANT: POSSANI, LOURIVAL
APPLICANT: CUEVAS-AGUIRRE, DELIA
APPLICANT: DONNER, PETER
APPLICANT: HECHLER, ULRICH
APPLICANT: HAENDLER, BERNARD
APPLICANT: HECHLER, ULRICH
TITLE OF INVENTION: THROMBIN INHIBITOR FROM THE SALIVA OF
TITLE OF INVENTION: PROSTOMIA
NUMBER OF SEQUENCES: 29
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/448,438
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/460,383
FILING DATE: 02-JUN-1995
APPLICATION NUMBER: WO PCT/DE93/91172
FILING DATE: 03-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE 4340798.6
FILING DATE: 25-NOV-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE 4328336.5
FILING DATE: 17-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE 4304731.9
FILING DATE: 12-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: DE 4241659.0
FILING DATE: 04-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: Hamlet-King, Diana
REGISTRATION/DOCKET NUMBER: 33,302
REFERENCE/DOCKET NUMBER: SCH 1474C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-243-6333
TELEFAX: 703-243-6410
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 160 amino acids
TYPE: amino acid sequence
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: SALIVA GLAND OF TRIATOMA PALIDIPENNIS
INDIVIDUAL ISOLATE: T1 5
FEATURE:
NAME/KEY: Protein
LOCATION: 19..160
US-08-448-438-8

Query Match 78.2%; Score 98.5; DB 2; Length 160;
Best Local Similarity 62.2%; Pred. No. 2.6e-08;
Matches 23; Conservative 0; Mismatches 1; Indels 13; Gaps 1;

QY 2 ISVSYNFALV-----DYVPERKSDTD 25
Db 117 ISVSYNFALVCRSITFTSQPKEDYLVLERTKSDTD 153

RESULT 9
US-09-489-039A-8097
Sequence 8097, Application US/09489039A
Patent No. 6610836
GENERAL INFORMATION:
APPLICANT: Gary Breton et. al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA

;; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
;; FILE REFERENCE: 2709.2004001
;; CURRENT APPLICATION NUMBER: US/09/489,039A
;; CURRENT FILING DATE: 2000-01-27
;; PRIOR APPLICATION NUMBER: US 60/117,747
;; PRIOR FILING DATE: 1999-01-29
;; NUMBER OF SEQ ID NOS: 14342
;; SEQ ID NO 8097
;; LENGTH: 548
;; TYPE: PRT
;; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-8097

Query Match 40.5%; Score 51; DB 4; Length 548;
Best Local Similarity 41.7%; Pred. No. 5.8;
Matches 10; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

Qy 2 ISVSYDNFALVDYLVFERTSDTD 25
Db 373 LNVGLDNFDLVKYLISQWQKDKD 396

RESULT 10
US-09-489-039A-13730
; Sequence 13730, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 13730
; LENGTH: 565
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-13730

Query Match 39.7%; Score 50; DB 4; Length 565;
Best Local Similarity 54.5%; Pred. No. 8.8;
Matches 12; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

Qy 4 VSYDNFALVDYLVFERTSDTD 25
Db 388 VGLDNFDLVKYLISQWQKDKD 409

RESULT 11
US-09-107-532A-4975
; Sequence 4975, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:

;; APPLICATION NUMBER: US/09/107,532A
;; FILING DATE: 30-Jun-1998
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 60/085,598
;; FILING DATE: 14 May 1998
;; APPLICATION NUMBER: 60/051571
;; FILING DATE: July 2, 1997
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Arinello, Pamela Denke
;; REGISTRATION NUMBER: 40,489
;; REFERENCE/DOCKET NUMBER: GTC-012
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (781)893-5007
;; TELEFAX: (781)893-8277
;; INFORMATION FOR SEQ ID NO: 4975:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 321 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; HYPOTHETICAL: YES
;; ORIGINAL SOURCE:
;; ORGANISM: Enterococcus faecium
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: (B) LOCATION 1...321
;; SEQUENCE DESCRIPTION: SEQ ID NO: 4975:
US-09-107-532A-4975

Query Match 37.7%; Score 47.5; DB 4; Length 321;
Best Local Similarity 50.0%; Pred. No. 11;
Matches 10; Conservative 3; Mismatches 4; Indels 3; Gaps 1;

Qy 6 YDNFALVDYLVFERTSDTD 25
Db 165 YDIVSLIDL---ETREDTD 181

RESULT 12
US-09-538-092-180
; Sequence 180, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Glot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuraPatSegFormatter Version 0.9
; SEQ ID NO 180
; LENGTH: 330
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (O)...(O)
; OTHER INFORMATION: Polypeptide Accession Number YDR267C
US-09-538-092-180

Query Match 37.3%; Score 47; DB 4; Length 330;
Best Local Similarity 57.1%; Pred. No. 14;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 2 ISVSYDNFALVDYL 15
Db 39 VSVKYDDFTLIDVL 52

RESULT 13
US-08-195-152-2
Sequence 2, Application US/08195152
Patent No. 5679541
GENERAL INFORMATION:
APPLICANT: Bonini, Nancy M.
APPLICANT: Leiserson, William M.
APPLICANT: Benzer, Seymour
TITLE OF INVENTION: PROGRAMMED CELL DEATH ANTAGONIST
TITLE OF INVENTION: PROTEINS
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: United States
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/195,152
FILING DATE: 14-FEB-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Treacartin, Richard F.
REGISTRATION NUMBER: 31,801
REFERENCE/DOCKET NUMBER: A-59551/RFT/RMS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 760 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-195-152-2

Query Match 36.9%; Score 46.5; DB 1; Length 760;
Best Local Similarity 47.6%; Pred. No. 46;
Matches 10; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 6 YDNFALVDYLVFERTKSD-TD 25
| | | | | : | | | | |
Db 8 YDNFSTLDYKVKRRPKTDHTD 28

RESULT 14
US-08-754-311B-2
Sequence 2, Application US/08754311B
Patent No. 6623937
GENERAL INFORMATION:
APPLICANT: Bonini, Nancy M.
APPLICANT: Leiserson, William M.
APPLICANT: Benzer, Seymour
TITLE OF INVENTION: PROGRAMMED CELL DEATH ANTAGONIST
TITLE OF INVENTION: PROTEINS
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: United States
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/754,311B
FILING DATE: 21-NOV-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/195,152
FILING DATE: 14-FEB-1994
ATTORNEY/AGENT INFORMATION:
NAME: Treacartin, Richard F.
REGISTRATION NUMBER: 31,801
REFERENCE/DOCKET NUMBER: A-59551/RFT/RMS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 760 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-754-311B-2

Query Match 36.9%; Score 46.5; DB 4; Length 760;
Best Local Similarity 47.6%; Pred. No. 46;
Matches 10; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 6 YDNFALVDYLVFERTKSD-TD 25
| | | | | : | | | | |
Db 8 YDNFSTLDYKVKRRPKTDHTD 28

RESULT 15
US-09-801-042-2
Sequence 2, Application US/09801042
Patent No. 6630332
GENERAL INFORMATION:
APPLICANT: RIEPING, MECHTILD
APPLICANT: THIERBACH, GEORG
APPLICANT: VAN DER REST, MICHEL EDUARD
APPLICANT: MOLENAAR, DOUWE
TITLE OF INVENTION: PROCESS FOR THE FERMENTATIVE PREPARATION OF L-THREONINE
FILE REFERENCE: MAS/21123/280410
CURRENT APPLICATION NUMBER: US/09/801,042
CURRENT FILING DATE: 2001-09-21
PRIOR APPLICATION NUMBER: 60/229,329
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: DE 100 34 833.5
PRIOR FILING DATE: 2000-07-18
PRIOR APPLICATION NUMBER: DE 101 03874.7
PRIOR FILING DATE: 2001-01-30
NUMBER OF SEQ ID NOS: 12
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 2
LENGTH: 529
TYPE: PRT
ORGANISM: Escherichia coli
US-09-801-042-2

Query Match 36.5%; Score 46; DB 4; Length 529;
Best Local Similarity 50.0%; Pred. No. 36;
Matches 11; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 4 VSYDNFALVDYLVFERTKSDPTD 25
| | | | | : | | | | |
Db 356 VGLDNFALVDYLVSYQVWLSSEED 377

RESULT 16
US-09-523-849-34

```
Sequence 34, Application US/09523849
; Patent No. 6458561
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Molteni, Angela
; APPLICANT: Magnaghi, Paola
; APPLICANT: Bobotti, Roberta
; APPLICANT: Scacheri, Emanuela
; APPLICANT: Isacchi, Antonella
; APPLICANT: Hodgson, Dave
; TITLE OF INVENTION: HUMAN NIM1 KINASE
; FILE REFERENCE: PC-0009 US
; CURRENT APPLICATION NUMBER: US/09/523,849
; CURRENT FILING DATE: 2000-03-13
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PERL Program
; SEQ ID NO 34
; LENGTH: 776
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Genbank Accession No. 6458561 95672676
US-09-523-849-34

Query Match          35.7%; Score 45; DB 4; Length 776;
Best Local Similarity 47.1%; Pred. No. 83;
Matches 8; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 5 SYDNFALVDYLVERTK 21
DB 330 SYNHPALYYLLERLR 346

RESULT 17
US-09-134-000C-5553
; Sequence 5553, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 5553
; LENGTH: 477
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-134-000C-5553

Query Match          35.3%; Score 44.5; DB 4; Length 477;
Best Local Similarity 52.9%; Pred. No. 56;
Matches 9; Conservative 3; Mismatches 4; Indels 1; Gaps 1;

QY 2 ISVSYNFALVDYLVPE 18
DB 200 IGMSYDNLIS-IDYTFE 215

RESULT 18
US-09-270-767-40476
; Sequence 40476, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
```

```
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 40476
; LENGTH: 154
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-40476

Query Match          34.9%; Score 44; DB 4; Length 154;
Best Local Similarity 50.0%; Pred. No. 17;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 2 ISVSYNFALVDYLVPE 17
DB 104 LSQALQHFQIVDYLVF 119

RESULT 19
US-09-270-767-55692
; Sequence 55692, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 55692
; LENGTH: 154
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-55692

Query Match          34.9%; Score 44; DB 4; Length 154;
Best Local Similarity 50.0%; Pred. No. 17;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 2 ISVSYNFALVDYLVPE 17
DB 104 LSQALQHFQIVDYLVF 119

RESULT 20
US-09-248-796A-15099
; Sequence 15099, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Kelch Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 15099
; LENGTH: 293
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-15099

Query Match          34.9%; Score 44; DB 4; Length 293;
Best Local Similarity 57.1%; Pred. No. 38;
Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 12 VDYLVEPTKSDTD 25
```


FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Arinello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 5418:
SEQUENCE CHARACTERISTICS:
LENGTH: 283 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc.feature
LOCATION: (B) LOCATION 1...283
SEQUENCE DESCRIPTION: SEQ ID NO: 5418:
US-09-107-532A-5418

Query Match 34.1%; Score 43; DB 4; Length 283;
Best Local Similarity 30.4%; Pred. No. 53;
Matches 7; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

OY 1 KISVSYDNFALVDYLVFERTKSD 23
DB 220 QMTKXYDNMTLIDWYSYSDHSD 242

RESULT 25
US-08-601-435-2
Sequence 2, Application US/08601435
Patent No. 5759801
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: DNA sequence coding for a protein
TITLE OF INVENTION: of A, thalana having a delta-5,7 sterol, delta-7
TITLE OF INVENTION: reductase activity, delta7-Red protein, production
TITLE OF INVENTION: process, strains of transformed yeasts, uses.
NUMBER OF SEQUENCES: 31
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
SOFTWARE:
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/601,435
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 9501723
FILING DATE: 15-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 9506517
FILING DATE: 01-JUN-1995
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 430 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-601-435-2

Query Match 34.1%; Score 43; DB 1; Length 430;

Best Local Similarity 41.7%; Pred. No. 87;
Matches 10; Conservative 3; Mismatches 7; Indels 4; Gaps 1;

OY 6 YDNFALVDY----LVFERTKSDTD 25
DB 381 FDNFALYFVTLILFDRARRDD 404

Search completed: May 9, 2005, 12:28:58
Job time : 44 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2005, 12:18:42 ; Search time 38 Seconds
(without alignments)
63.301 Million cell updates/sec

Title: US-10-694-847-3

Perfect score: 126
Sequence: 1 KISVSYDNALVDLVFERTKSDTD 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : PIR 769:***
1: p1r1:***
2: p1r2:***
3: p1r3:***
4: p1r4:***

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	49	38.9	140	2 T01027	hypothetical prote
2	49	38.9	317	2 P95913	probable polysacch
3	48	38.1	481	2 I49072	protein kinase - m
4	48	38.1	798	2 JC7500	glk protein - chic
5	47	37.3	330	2 S70127	hypothetical prote
6	46.5	36.9	760	2 A45174	eye cell developme
7	46	36.5	405	2 AB1461	B. subtilis yads p
8	46	36.5	408	2 AC1098	B. subtilis yads p
9	46	36.5	548	2 E85860	hypothetical prote
10	46	36.5	548	2 H64990	probable malate de
11	46	36.5	548	2 C91016	probable malate de
12	46	36.5	562	2 A82743	probable malate de
13	46	36.5	758	2 I64084	hemoglobin recepto
14	45.5	36.1	158	2 S73833	MG236 homolog P10
15	45.5	36.1	674	2 G72204	hypothetical prote
16	45	35.7	383	2 B87579	hypothetical prote
17	45	35.7	1116	2 T38073	serine/chreonine-p
18	44.5	35.3	515	2 S63382	hypothetical prote
19	44.5	35.3	1249	2 AC1065	helicase related p
20	44.5	35.3	4092	1 S38128	dynein heavy chain
21	44	34.9	331	2 E83819	mga-like (mycopla
22	44	34.9	331	2 E83819	hypothetical prote
23	44	34.9	434	2 A42512	GSR protein - vacc
24	44	34.9	434	2 T37350	probable 49.8K pro
25	44	34.9	434	2 A72159	ISR protein - vari
26	44	34.9	434	2 S33082	GSR protein - vari
27	44	34.9	434	2 T28505	hypothetical prote
28	44	34.9	525	2 AC0452	maltose transport
29	44	34.9	595	2 C97458	maltose-quinone oxi

30	44	34.9	595	2 AD2676	maltose:quinone oxi
31	44	34.9	1026	2 T39612	hypothetical prote
32	43.5	34.5	230	2 T41660	glutathione S-tran
33	43	34.1	185	2 C90548	lipoprotein [impor
34	43	34.1	215	2 H69251	hypothetical prote
35	43	34.1	222	1 KAB052	alpha-s2-casein pr
36	43	34.1	293	2 G86901	hypothetical prote
37	43	34.1	293	2 G86901	acetylglutamate ki
38	43	34.1	395	2 T50392	hypothetical coile
39	42.5	33.7	169	2 B72345	holo-(acyl) carrier
40	42.5	33.7	336	2 A72803	minor tail subunit
41	42.5	33.7	476	2 G90512	p60-like (mycoplas
42	42	33.3	213	1 XEBCF	chloramphenicol O-
43	42	33.3	215	2 E97937	hypothetical prote
44	42	33.3	228	2 T11063	hypothetical prote
45	42	33.3	283	2 T11063	NADH2 dehydrogenas
46	42	33.3	333	2 H89015	protein B0213.7 [1
47	42	33.3	375	2 T03325	gene 118 protein -
48	42	33.3	402	2 B90322	glycosyltransferas
49	42	33.3	432	2 F96540	sterol delta7 redu
50	42	33.3	472	2 T43084	transfer complex p
51	42	33.3	474	2 AE3426	mannose-1-phosphat
52	42	33.3	649	2 H72283	beta-galactosidase
53	42	33.3	792	2 H83736	subtilisin-type al
54	42	33.3	889	2 A35679	rep protein - slim
55	42	33.3	1170	1 TSHUP1	thrombospondin 1 p
56	42	33.3	1170	2 A40558	thrombospondin 1 p
57	42	33.3	1178	1 A39804	thrombospondin pre
58	41.5	32.9	214	4 S58320	hypothetical prote
59	41.5	32.9	325	2 C83158	probable 2-hydroxy
60	41.5	32.9	674	2 H64613	flagellar hook-ass
61	41.5	32.9	675	2 S25005	dnak-type molecula
62	41.5	32.9	675	2 S19140	dnak-type molecula
63	41.5	32.9	685	2 A71901	probable flagellar
64	41.5	32.9	777	2 T25761	hypothetical prote
65	41	32.5	121	2 A11646	hypothetical prote
66	41	32.5	194	2 A11869	hypothetical prote
67	41	32.5	215	2 G95069	ABC transporter, A
68	41	32.5	293	2 D90153	cobalamin biosynth
69	41	32.5	440	2 T41766	ARIR-1 orf20/21 -
70	41	32.5	488	2 A81250	hypothetical prote
71	41	32.5	758	2 T25751	hypothetical prote
72	41	32.5	949	2 T06003	metal-transporting
73	41	32.5	1434	2 C82923	DNA-directed RNA p
74	41	32.5	1660	2 A84647	hypothetical prote
75	40.5	32.1	375	2 T24903	hypothetical prote
76	40.5	32.1	1066	2 G84746	hypothetical prote
77	40	31.7	81	2 D97063	uncharacterized pr
78	40	31.7	142	2 D72040	conserved hypotet
79	40	31.7	142	2 C86583	CT635 hypothetical
80	40	31.7	166	2 T25773	hypothetical prote
81	40	31.7	250	2 AH1224	cobalamin biosynth
82	40	31.7	250	2 AB1578	cobalamin biosynth
83	40	31.7	255	1 VG8E2E	glycoprotein E - e
84	40	31.7	264	2 D71197	hypothetical prote
85	40	31.7	275	2 S77476	tRNA-pseudouridine
86	40	31.7	295	2 T33681	hypothetical prote
87	40	31.7	305	2 B72865	clostridin-relate
88	40	31.7	341	2 T33990	hypothetical prote
89	40	31.7	353	2 G71910	hypothetical prote
90	40	31.7	360	2 D95185	proline dipeptidas
91	40	31.7	360	2 D98052	X-Pro dipeptidase
92	40	31.7	383	1 H64601	conserved hypotet
93	40	31.7	406	1 AG0548	probable ABC-trans
94	40	31.7	496	2 S75790	hypothetical prote
95	40	31.7	514	1 MMECMF	maltose transport
96	40	31.7	514	2 AC1014	inner membrane pro
97	40	31.7	514	2 S05332	transport system p
98	40	31.7	514	2 H91255	part of maltose pe
99	40	31.7	514	2 D86096	inner membrane pro
100	40	31.7	514	2 S20604	inner membrane pro

ALIGNMENTS

RESULT 1

T01027

hypothetical protein YUP8H12R.10 - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 05-Feb-1999 #sequence_revision 05-Feb-1999 #text_change 09-Jul-2004

C/Accession: T01027

R/Theologis, A.; Vlasakakaia, V.S.; Osborne, B.I.; Schwartz, J.R.; Federspiel, N.A.; Kwar

Oefner, P.; Davis, R.W.

submitted to the EMBL Data Library, May 1998

A/Description: Arabidopsis thaliana chromosome 1 YAC YUP8H12R sequence.

A/Reference number: 214227

A/Accession: T01027

A/Status: translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1140 <THE>

A/Cross-references: UNIPROT:Q64523; EMBL:AC002986; NID:G2494106; PID:G3152578; GSPDB:GNC

C/Genetics:

A/Map position: 1

A/Intons: 58/3; 72/2; 105/2; 121/2; 132/3

C/Superfamily: Arabidopsis thaliana hypothetical protein YUP8H12R.10

Query Match 38.9%; Score 49; DB 2; Length 140;

Best Local Similarity 43.8%; Pred. No. 2.3;

Matches 7; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 4 VSYDNFALVDYLVFERK 19

DB 98 LAYDHFLDLDFLIMRR 113

RESULT 2

P95913 probable polyaccharide export-associated protein [imported] - Sinorhizobium meliloti (st

C/Species: Sinorhizobium meliloti

C/Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004

C/Accession: P95913

R/Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan

Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001

A/Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endo

A/Reference number: A95842; MUID:21396508; PMID:11481431

A/Accession: P95913

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-317 <KUR>

A/Cross-references: UNIPROT:Q92VX3; GB:AL591985; PIDN:CAC48974.1; PID:G15140459; GSPDB:G

A/Experimental source: strain 1021, megaplasmid pSymB

R/Galbert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,

peta, D.; Chain, P.; Cowie, A.; Davis, R.W.; Drenno, S.; Federspiel, N.A.; Fisher, R.F.;

L.; Hyman, R.W.; Jones, T.

Science 293, 668-672, 2001

A/Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,

hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.

A/Title: The composite genome of the legume symbiont Sinorhizobium meliloti.

A/Reference number: A96039; MUID:21368234; PMID:11474104

A/Contents: annotation

C/Genetics:

A/Genes: kPR/kpse; SMB20831

A/Genome: plasmid

Query Match 38.9%; Score 49; DB 2; Length 317;

Best Local Similarity 42.1%; Pred. No. 5.5;

Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 7 DNFALVDYLVFERKSPD 25

DB 13 DSYLMDYILSERMADAD 31

RESULT 3

149072 protein kinase - mouse (fragment)

C/Species: Mus musculus (house mouse)

C/Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 16-Aug-2004

C/Accession: 149072

R/Ruiz, J.C.; Conlon, F.L.; Robertson, E.J.

Mech. Dev. 48, 153-164, 1994

A/Title: Identification of novel protein kinases expressed in the myocardium of the deve

A/Reference number: 149071; MUID:95200798; PMID:7893599

A/Accession: 149072

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-481 <RES>

A/Cross-references: EMBL:U11494; NID:G595420; PIDN:AAA67926.1; PID:G595421

C/Superfamily: protein kinase homology

C/Keywords: ATP

P/71-324/Domain: protein kinase homology <KIN>

P/79-87/Region: protein kinase ATP-binding motif

Query Match 38.1%; Score 48; DB 2; Length 481;

Best Local Similarity 52.9%; Pred. No. 12;

Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 5 SYDNFALVDYLVFERK 21

DB 376 SYNHFAIYLLERLK 392

RESULT 4

JC7500 qik protein - chicken

N/Alternate names: Qln-induced kinase

C/Species: Gallus gallus (chicken)

C/Date: 17-Nov-2000 #sequence_revision 17-Nov-2000 #text_change 09-Jul-2004

C/Accession: JC7500

R/Xia, Y.; Zhang, Z.; Kruse, U.; Vogt, P.K.; Li, J.

Biochem. Biophys. Res. Commun. 276, 564-570, 2000

A/Title: The new serine-threonine kinase, Qik, is a target of the qin oncogene.

A/Reference number: JC7500

A/Contents: Embryo fibroblasts

A/Accession: JC7500

A/Molecule type: mRNA

A/Residues: 1-798 <XIA>

A/Cross-references: UNIPROT:Q91A88; GB:AF219232

C/Comment: This protein, a member of the AMPK/SNF1 family of serine/threonine kinases, fr

C/Genetics:

A/Genes: qik

C/Keywords: protein kinase

Query Match 38.1%; Score 48; DB 2; Length 798;

Best Local Similarity 52.9%; Pred. No. 21;

Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 5 SYDNFALVDYLVFERK 21

DB 329 SYNHFAIYLLERLK 345

RESULT 5

S70127 hypothetical protein YDR267c - yeast (Saccharomyces cerevisiae)

N/Alternate names: hypothetical protein D9954.11

C/Species: Saccharomyces cerevisiae

C/Date: 24-Aug-1996 #sequence_revision 06-Sep-1996 #text_change 16-Aug-2004

C/Accession: S70127

R/Le, T.

submitted to the EMBL Data Library, May 1996

A/Description: The sequence of S. cerevisiae cosmid 9954.

A/Reference number: S70124

A/Accession: S70127

A/Molecule type: DNA

A/Residues: 1-330 <LET>

A/Cross-references: UNIPROT:Q05583; EMBL:U51030; NID:G1332633; PID:G1230640; GSPDB:GN000

C/Genetics:
 A:Gene: MIPS:YDR267c
 A:Cross-references: SGD:S0002675
 A:Map position: 4R
 C:Superfamily: WD repeat homology
 F:54-87/Domain: WD repeat homology <MD1>
 F:103-136/Domain: WD repeat homology <MD2>
 F:149-182/Domain: WD repeat homology <MD3>

Query Match 37.3%; Score 47; DB 2; Length 330;
 Best Local Similarity 57.1%; Pred. No. 12;
 Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 ISVYNFALVDYL 15
 DB 39 VSVKXDDFTLIDVL 52

RESULT 6

A45174

eye cell development gene *eya* protein, splice form 1 - fruit fly (*Drosophila melanogaster*)
 C/Species: *Drosophila melanogaster*
 C/Date: 30-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
 C/Accession: A45174

R:Bouini, N.M.; Leiserson, W.M.; Benzer, S.

Cell 72, 379-395, 1993

A>Title: The *eya* gene: genetic control of cell survival and differentiation in *Drosophila*

A:Reference number: A45174; MUID:93161413; PMID:8431945

A:Accession: A45174

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-760 <BON>

A:Cross-references: UNIPROT:Q05201; GB:L08501; NID:g157975; PID:g157976

A:Experimental source: adult head

A:Note: sequence extracted from NCBI backbone (NCBIP:124850)

C/Genetics:

A:Gene: FlyBase:cli; *eya*; eyes absent

A:Cross-references: FlyBase:Fbgn0000320

Query Match 36.9%; Score 46.5; DB 2; Length 760;
 Best Local Similarity 47.6%; Pred. No. 35;
 Matches 10; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 6 YDNFALVDYLFERTKSD 25
 DB 8 YQNFSTLDYKXKRPKTDHT 28

RESULT 7

AB1461

B. subtilis *yabE* protein homolog lin0225 [imported] - *Listeria innocua* (strain Clp11262)
 C/Species: *Listeria innocua*
 C/Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
 C/Accession: AB1461

R:Jlares, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Duesurget, O.; Entian, K.D.; Fsihi, H.

Science 294, 849-852, 2001

A:Authors: Kreft, U.; Kuhn, M.; Kunst, F.; Kurapkak, G.; Madueno, E.; Maltournam, A.; Matok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of *Listeria* species.

A:Reference number: AB1077; MUID:21537279; PMID:11679669

A:Accession: AB1461

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-405 <GLA>

A:Cross-references: UNIPROT:Q92F81; GB:AL592022; PIDN:CA95458.1; PID:g16412644; GSPDB:G

A:Experimental source: strain Clp11262

C/Genetics:

A:Gene: lin0225

Query Match 36.5%; Score 46; DB 2; Length 405;
 Best Local Similarity 71.4%; Pred. No. 22;

Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 10 ALVDYLVFERTKSD 23
 DB 29 ALVFVFEKTKND 42

RESULT 8

AC1098

B. subtilis *yabE* protein homolog lmo0186 [imported] - *Listeria monocytogenes* (strain EGD-
 C/Species: *Listeria monocytogenes*
 C/Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
 C/Accession: AC1098

R:Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Duesurget, O.; Entian, K.D.; Fsihi, H.

Science 294, 849-852, 2001

A:Authors: Kreft, U.; Kuhn, M.; Kunst, F.; Kurapkak, G.; Madueno, E.; Maltournam, A.; Matok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of *Listeria* species.

A:Reference number: AB1077; MUID:21537279; PMID:11679669

A:Accession: AC1098

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-408 <GLA>

A:Cross-references: UNIPROT:Q9YAE4; GB:NC_003210; PIDN:CA98401.1; PID:g16409543; GSPDB:G

A:Experimental source: strain EGD-e

C/Genetics:

A:Gene: lmo0186

Query Match 36.5%; Score 46; DB 2; Length 408;
 Best Local Similarity 71.4%; Pred. No. 22;
 Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 10 ALVDYLVFERTKSD 23
 DB 29 ALVFVFEKTKND 42

RESULT 9

E85860

hypothetical protein yojH [imported] - *Escherichia coli* (strain O157:H7, substrain EDL93)
 C/Species: *Escherichia coli*
 C/Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 12-Jun-2003
 C/Accession: E85860

R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew, M.; Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca, Nature 409, 529-533, 2001

A>Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.

A:Reference number: AB5480; MUID:21074935; PMID:11206551

A:Accession: E85860

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-548 <STO>

A:Cross-references: GB:AE005174; NID:g12516545; PIDN:AA657345.1; GSPDB:GN00145; UMGD:2346

A:Experimental source: strain O157:H7, substrain EDL933

C/Genetics:

A:Gene: yojH

C:Superfamily: malate dehydrogenase (acceptor)

Query Match 36.5%; Score 46; DB 2; Length 548;
 Best Local Similarity 50.0%; Pred. No. 30;
 Matches 11; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 4 VSYDNFALVDYLVFERTKSD 25
 DB 375 VGDNDLVKYLVSQVWLSEED 396

RESULT 10

H64990

probable malate dehydrogenase (acceptor) (EC 1.1.99.16) yojH [similarity] - *Escherichia coli*
 C/Species: *Escherichia coli*

C>Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004
 C/Accession: H64990
 R/Author: F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Col
 A.; Rose, D.J.; Mau, B.; Shao, Y.
 Science 277, 1453-1462, 1997
 A/Title: The complete genome sequence of *Escherichia coli* K-12.
 A/Reference number: A64720; MUID:97426617; PMID:9278503
 A/Accession: H64990
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1548 <BLAT>
 A/Cross-references: UNIPROT:P33940; GB:AE000310; GE:U00096; NID:92367131; PIDN:AACT5270
 A/Experimental source: strain K-12, substrain MG1655
 C/Genetics:
 A:Gene: yojH
 C:Superfamily: malate dehydrogenase (acceptor)
 C/Keywords: oxidoreductase

Query Match 36.5%; Score 46; DB 2; Length 548;
 Best Local Similarity 50.0%; Pred. No. 30;
 Matches 11; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 4 VSYDNFALVDVLYPERKSDTD 25
 DB 375 VGLDNFDLVKLYVSQVWLSEED 396

RESULT 11

C91016
 Probable malate quione oxidoreductase EC63099 [imported] - *Escherichia coli* (strain O15
 C/Species: *Escherichia coli*
 C/Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
 C/Accession: C91016
 R/Author: T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
 gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
 DNA Res. 8, 11-22, 2001
 A/Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and gene
 A/Reference number: A9629; MUID:21156231; PMID:11258796
 A/Accession: C91016
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1548 <HAY>
 A/Cross-references: UNIPROT:Q8XEA5; GB:BA000007; PIDN:BA836522.1; PID:G13362568; GSPDB:C
 A/Experimental source: strain O157:H7, substrain RIMD 0509952
 C/Genetics:
 A:Gene: EC63099
 C:Superfamily: malate dehydrogenase (acceptor)

Query Match 36.5%; Score 46; DB 2; Length 548;
 Best Local Similarity 50.0%; Pred. No. 30;
 Matches 11; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 4 VSYDNFALVDVLYPERKSDTD 25
 DB 375 VGLDNFDLVKLYVSQVWLSEED 396

RESULT 12

A82743
 Probable malate dehydrogenase (acceptor) (EC 1.1.99.16) XF0942 [imported] - *Xylella fast*
 C/Species: *Xylella fastidiosa*
 C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
 C/Accession: A82743
 R/Author: The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
 Nature 406, 151-157, 2000
 A/Title: The genome sequence of the plant pathogen *Xylella fastidiosa*.
 A/Reference number: A82515; MUID:20365717; PMID:10910347
 A/Note: for a complete list of authors see reference number A59328 below
 A/Accession: A82743
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1562 <SIM>
 A/Cross-references: UNIPROT:Q9PET6; GB:AE003933; GB:AE003849; NID:9105863; PIDN:AAF8375

A/Experimental source: strain 9a5c
 R/Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
 Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carre, H.
 as-Neto, E.; Docena, C.; El-Dorri, H.; Facinelli, A.P.; Ferreira, A.J.S.
 submitted to Genbank, June 2000

A/Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohme
 J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.B.; Laigre
 Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E.
 A/Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
 F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A.
 Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Sencelli, R.V.; Sawaaki
 A/Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
 M.; Tshahko, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Ze
 A/Reference number: A59328
 A/Contents: annotation
 C/Genetics:
 A:Gene: XF0942
 C:Superfamily: malate dehydrogenase (acceptor)
 C/Keywords: oxidoreductase

Query Match 36.5%; Score 46; DB 2; Length 562;
 Best Local Similarity 41.7%; Pred. No. 31;
 Matches 10; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 2 ISVSYDNFALVDVLYPERKSDTD 25
 DB 382 LSVGNKNNIDLVKLYMKQATLTDD 405

RESULT 13

I64084
 hemoglobin receptor homolog - *Haemophilus influenzae* (strain Rd KW20)
 C/Species: *Haemophilus influenzae*
 C/Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 15-Mar-2004
 C/Accession: I64084
 R/Author: R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirnesh, E.F.; Kertavage, A.
 R.; Gecayene, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodok, A.; Kelley, J.M.; Weidman, J.
 D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghegan, N.S.M.
 Science 269, 496-512, 1995
 A/Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,
 A/Title: Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd.
 A/Reference number: A64000; MUID:95350630; PMID:7542800
 A/Accession: I64084
 A/Status: nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1758 <TIGR>
 A/Cross-references: GB:U32749; GB:I42023; NID:91573658; PID:G1573661; TIGR:HI0661
 C/Genetics:
 A:Start codon: GTG
 F/385-758/Domain: lonB-dependent receptor carboxyl-terminal homology <TNC>

Query Match 36.5%; Score 46; DB 2; Length 758;
 Best Local Similarity 53.8%; Pred. No. 42;
 Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 6 YDNFALVDVLYPER 18
 DB 387 FDNFVITDYLSPD 399

RESULT 14

S73833
 MG236 homolog F10 orf158 - *Mycoplasma pneumoniae* (strain ATCC 29342)
 C/Species: *Mycoplasma pneumoniae*
 A/Variety: ATCC 29342
 C/Date: 27-Feb-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
 R/Himmelsbach, R.; Hilbert, H.; Plagens, H.; Pirk, E.; Li, B.C.; Herrmann, R.
 Nucleic Acids Res. 24, 4420-4449, 1996
 A/Title: Complete sequence analysis of the genome of the bacterium *Mycoplasma pneumoniae*.
 A/Reference number: S73327; MUID:97105885; PMID:8948633
 A/Accession: S73833
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown

Matches	7; Conservative	5; Mismatches	6; Indels	0; Gaps	0; N/A
QY	8 NFALVDYLVFERTKSDTD	25			
	:: :::				
Db.	294 NMGMVDLYLNYQNMADTD	311			

RESULT 23

G5R:protein vaccinia virus (strain Copenhagen)
C:Species: vaccinia virus
A:Note: host Homo sapiens (man)
C:Date: 09-Nov-1990 #revision 09-Nov-1990 #text_change 09-Jul-2004
C:Accession: A42512
R:Johnson, G.P.
submitted to GenBank, June 1990
A:Reference number: A33172
A:Accession: A42512
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-434 <JOH>
A:Cross-references: UNIPROT:P21026
C:Superfamily: vaccinia virus probable 49.8K protein

Query Match

Query Match	34.9%	Score 44;	DB 2;	Length 434;
Best Local Similarity	60.0%;	Pred. No. 48;		
Matches 9;	Conservative 1;	Mismatches 5;	Indels 0;	Gaps 0

```
QY      11  LVDYLVFERTKSDTD 25
          | : | | | | |
Db      363 LSSYIDFENTKSDID 377
```

RESULT 24

Probable 49.8K protein - vaccinia virus (strain Ankara)
 C:Species: vaccinia virus
 A:Variety: strain Ankara
 C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 09-Jul-2004
 C:Accession: J37350
 R:Annotator, G.; Scheiflinger, F.; Falkner, F.G.; Dörner, F.
 submitted to the EMBL Data Library, March 1997
 A:Description: The complete genomic sequence of the Modified Vaccinia Ankara (MVA) strain
 A:Reference number: J20877
 A:Accession: J37350
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-434 <ANT>
 A:Cross-references: UNIPROT:O57196; EMBL:U94848; PIDDN:AA96493.1
 A:Experimental source: strain Ankara
 C:Genetics:
 A:Note: MVA074R
 A:Superfamily: vaccinia virus probable 49.8K protein

Query Match

Query Match	34.9%	Score 44	DB 2	Length 434
Best Local Similarity	60.0%	Pred. No. 48		
Matches 9, Conservative		1	Mismatches 5	Indels 0
			Gaps	0

```
Qy      11  LVDYLVFERTKSDTD 25
          |  |  |  |  |  |  |
Db.     363 LSSYIDFENTKSDID 377
```

RESULT 25

I:ISR protein - variola minor virus (strain Garcia-1966)
 C:Species: variola minor virus
 C:Date: 24-Nov-1999 #sequence_revision 24-Nov-1999 #text_change 09-Jul-2004
 C:Accession: A72150
 R:Shchelkunov, S.N.; Totmenin, A.V.; Gultarov, V.V.; Safonov, P.F.; Masang, R.F.; Logaev
 submitted to Genbank, March 1998
 A:Description: Analysis of the complete coding sequence of DNA of atastrim variola minor
 ;Reference number: A72150

```

A:Accession: A72159
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-434 <SHC>
A:Cross-references: UNIPROT:Q89210; GB:Y16780; NID:g5830555; PIDN:CAB54667.1; PID:g58306.
A:Experimental source: strain Garcia-1966
C:Genetics:
A:Gene: 15R
C:Superfamily: vaccinia virus probable 49.8K protein

Query Match          34.9%; Score 44; DB 2; Length 434;
Best Local Similarity 60.0%; Pred. No. 48;
Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Oy      11  LVDYLVPERTKSDTD 25
      |  |  |  |  |  |  |
Db      363  LSSYIDFENTKSDID 377

```

Search complete

Search completed: May 9, 2005, 12:28:12
Job time : 42 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2005, 12:21:48 ; Search time 131 Seconds
(without alignments)
63.664 Million cell updates/sec

Title: US-10-694-847-3

Perfect score: 126
Sequence: 1 KISVSVDNFALVDYLVFERTKSDPT 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1428581 seqs, 333598853 residues

Total number of hits satisfying chosen parameters: 1428581

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 100 summaries

Database :

Published Applications AA:*

1: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubppaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubppaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubppaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubppaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubppaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubppaa/US09C_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubppaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubppaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubppaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubppaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/2/pubppaa/US10_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubppaa/US11_NEW_PUB.pep:*
19: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB.pep:*
20: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	51	40.5	544	15	US-10-282-122A-59516
2	50	39.7	549	15	US-10-282-122A-56377
3	48	38.1	779	10	US-09-823-187-94
4	48	38.1	779	15	US-10-231-913-121
5	48	38.1	782	10	US-09-823-187-26
6	48	38.1	783	9	US-09-815-915-2
7	48	38.1	783	10	US-09-823-187-90
8	48	38.1	783	14	US-10-393-316-2
9	48	38.1	786	10	US-09-823-187-91
10	48	38.1	786	15	US-10-231-913-118
11	48	38.1	798	10	US-09-823-187-95
12	48	38.1	798	15	US-10-231-913-122
13	48	38.1	826	15	US-10-425-114-54182

14	48	38.1	1193	16	US-10-437-963-201611	Sequence 201611, A
15	47	37.3	511	9	US-09-764-868-834	Sequence 834, App
16	47	37.3	522	11	US-09-764-875-636	Sequence 636, App
17	47	37.3	589	15	US-10-425-114-39113	Sequence 39113, A
18	47	37.3	896	16	US-10-408-765A-1811	Sequence 1811, App
19	47	37.3	923	15	US-10-258-106-7	Sequence 7, Appl
20	47	37.3	1673	8	US-10-437-963-149049	Sequence 149049, A
21	46.5	36.9	760	16	US-08-754-111B-2	Sequence 2, Appl
22	46	36.5	77	16	US-10-437-963-127926	Sequence 127926, A
23	46	36.5	529	9	US-09-801-042-2	Sequence 2, Appl
24	46	36.5	538	9	US-09-815-242-10223	Sequence 10223, A
25	46	36.5	548	15	US-10-282-122A-56616	Sequence 56616, A
26	45	35.7	776	10	US-09-823-187-92	Sequence 92, Appl
27	45	35.7	776	10	US-09-823-187-93	Sequence 93, Appl
28	45	35.7	776	14	US-10-195-101-34	Sequence 34, Appl
29	45	35.7	776	15	US-10-231-913-119	Sequence 119, App
30	45	35.7	776	15	US-10-231-913-120	Sequence 120, App
31	45	35.7	1116	15	US-10-369-493-2069	Sequence 2069, App
32	44.5	35.3	264	16	US-10-437-963-126279	Sequence 126279, A
33	44.5	35.3	363	16	US-10-437-963-126280	Sequence 126280, A
34	44.5	35.3	1249	15	US-10-282-122A-76279	Sequence 76279, A
35	44.5	35.3	4092	15	US-10-369-493-22278	Sequence 22278, A
36	44	34.9	171	16	US-10-437-963-159018	Sequence 159018, A
37	44	34.9	224	9	US-09-823-847-4	Sequence 4, Appl
38	44	34.9	653	15	US-10-389-566-805	Sequence 805, App
39	44	34.9	698	15	US-10-389-566-694	Sequence 694, App
40	44	34.9	912	15	US-10-389-566-693	Sequence 693, App
41	44	34.9	1293	16	US-10-437-963-186644	Sequence 186644, A
42	43.5	34.5	349	15	US-10-282-122A-72273	Sequence 72273, A
43	43	34.1	65	15	US-10-424-599-281106	Sequence 281106, A
44	43	34.1	66	9	US-09-922-199A-4	Sequence 4, Appl
45	43	34.1	114	16	US-10-437-963-125803	Sequence 125803, A
46	43	34.1	171	15	US-10-424-599-2164074	Sequence 216407, A
47	43	34.1	172	15	US-10-424-599-164074	Sequence 164074, A
48	43	34.1	172	15	US-10-425-114-62404	Sequence 62404, A
49	43	34.1	233	15	US-10-369-493-1111	Sequence 1111, App
50	43	34.1	425	15	US-10-402-84-25	Sequence 25, Appl
51	43	34.1	400	15	US-10-224-880C-31	Sequence 31, Appl
52	43	34.1	509	14	US-10-156-761-7900	Sequence 7900, App
53	43	34.1	963	15	US-10-282-122A-61097	Sequence 61097, A
54	43	34.1	1779	16	US-10-437-963-155515	Sequence 155515, A
55	42.5	33.7	128	15	US-10-424-599-219065	Sequence 219065, A
56	42.5	33.7	131	10	US-09-746-783-192	Sequence 192, App
57	42.5	33.7	132	9	US-09-984-245-183	Sequence 183, App
58	42.5	33.7	132	10	US-09-966-262-183	Sequence 183, App
59	42.5	33.7	132	10	US-09-983-966-183	Sequence 183, App
60	42.5	33.7	132	14	US-10-059-395-183	Sequence 183, App
61	42.5	33.7	132	14	US-10-143-090-183	Sequence 183, App
62	42.5	33.7	132	17	US-10-960-251-183	Sequence 183, App
63	42.5	33.7	139	15	US-10-296-115-1473	Sequence 1473, App
64	42.5	33.7	159	9	US-09-771-383-6	Sequence 6, Appl
65	42.5	33.7	159	10	US-09-770-834-7	Sequence 7, Appl
66	42.5	33.7	159	15	US-10-717-138-7	Sequence 7, Appl
67	42.5	33.7	182	15	US-10-276-774-2296	Sequence 2296, App
68	42.5	33.7	183	15	US-10-262-511-132	Sequence 132, App
69	42	33.3	31	9	US-09-155-203-3	Sequence 3, Appl
70	42	33.3	38	9	US-09-155-203-2	Sequence 2, Appl
71	42	33.3	39	9	US-09-155-203-1	Sequence 1, Appl
72	42	33.3	71	16	US-10-437-963-108332	Sequence 108332, A
73	42	33.3	98	15	US-10-424-599-284693	Sequence 284693, A
74	42	33.3	177	10	US-09-956-622A-33	Sequence 33, Appl
75	42	33.3	224	16	US-10-437-963-143174	Sequence 143174, A
76	42	33.3	226	16	US-10-767-701-37515	Sequence 37515, A
77	42	33.3	227	15	US-10-719-466-53	Sequence 53, Appl
78	42	33.3	227	17	US-10-782-968-53	Sequence 53, Appl
79	42	33.3	406	15	US-10-425-114-70484	Sequence 70484, A
80	42	33.3	408	9	US-09-817-774-41	Sequence 41, Appl
81	42	33.3	411	9	US-09-817-774-43	Sequence 43, Appl
82	42	33.3	422	9	US-09-817-774-29	Sequence 29, Appl
83	42	33.3	422	9	US-09-817-774-39	Sequence 39, Appl
84	42	33.3	432	9	US-09-817-774-45	Sequence 45, Appl
85	42	33.3	455	16	US-10-437-963-134006	Sequence 134006, A
86	42	33.3	476	9	US-09-817-774-31	Sequence 31, Appl

87 42 33.3 494 15 US-10-425-114-40446 Sequence 40446, A
88 42 33.3 526 15 US-10-282-122A-48057 Sequence 48057, A
89 42 33.3 608 15 US-10-437-963-171608 Sequence 171608, A
90 42 33.3 648 15 US-10-425-114-71724 Sequence 71724, A
91 42 33.3 649 15 US-10-369-493-3010 Sequence 3010, A
92 42 33.3 664 14 US-10-319-762-10 Sequence 10, Appl
93 42 33.3 665 14 US-10-319-762-6 Sequence 6, Appl
94 42 33.3 665 14 US-10-424-599-229557 Sequence 229557, A
95 42 33.3 723 14 US-10-168-424-23 Sequence 23, Appl
96 42 33.3 1152 9 US-09-919-603-1 Sequence 1, Appl
97 42 33.3 1170 14 US-10-020-141-12 Sequence 12, Appl
98 42 33.3 1170 14 US-10-017-721-2 Sequence 2, Appl
99 42 33.3 1170 14 US-10-021-660-114 Sequence 11, Appl
100 42 33.3 1170 14 US-10-008-093-2 Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-10-282-122A-59516
; Sequence 59516, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zykkind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 59516
; LENGTH: 544
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-10-282-122A-59516

Query Match 40.5%; Score 51; DB 15; Length 544;
Best Local Similarity 41.7%; Pred. No. 28;
Matches 10; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 2 ISVSYNFALVDYLVFERKSDTD 25
Db 369 LNVGLDNFDLVKYLISQVMSDD 392

RESULT 2
US-10-282-122A-56377
; Sequence 56377, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zykkind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 56377
; LENGTH: 549
; TYPE: PRT
; ORGANISM: Enterobacter cloacae
US-10-282-122A-56377

Query Match 39.7%; Score 50; DB 15; Length 549;
Best Local Similarity 45.8%; Pred. No. 40;
Matches 11; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 2 ISVSYNFALVDYLVFERKSDTD 25
Db 374 MDVGLDNFDLVKYLISQVMSDD 397

RESULT 3
US-09-823-187-94
; Sequence 94, Application US/09823187
; Publication No. US20030096952A1
; GENERAL INFORMATION:
; APPLICANT: Burgess, Catherine
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Liu, Xiaohong
; APPLICANT: Majumder, Kumud

```

; APPLICANT: Padigar, Muralidhar
; APPLICANT: Patturajan, Meera
; APPLICANT: Shinkete, Richard A
; APPLICANT: Spaderna, Steven K
; APPLICANT: Spytek, Kimberly
; APPLICANT: Taupier, Raymond J
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-745
; CURRENT APPLICATION NUMBER: US/09/823,187
; CURRENT FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/193,339
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 60/193,205
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 60/195,343
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: 60/195,088
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,005
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,792
; PRIOR FILING DATE: 2000-04-10
; PRIOR APPLICATION NUMBER: 60/196,556
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: 60/197,081
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: 60/197,525
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/197,087
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 103
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 94
; LENGTH: 779
; TYPE: PRT
; ORGANISM: Mus musculus
; US-09-823-187-94

Query Match      38.1%; Score 48; DB 10; Length 779;
Best Local Similarity 52.9%; Pred. No. 1.2e+02;
Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY      5 SYDNFALVDYLVFERTK 21
DB      330 SYNHPALYYLLERLK 346

RESULT 4
US-10-231-913-121
; Sequence 121, Application US/10231913
; Publication No. US20040005576A1
; GENERAL INFORMATION:
; APPLICANT: Guo, Xiaojia S.
; APPLICANT: Li, Li
; APPLICANT: Patturajan, Meera
; APPLICANT: Shinkete, Richard A.
; APPLICANT: Casman, Stacie J.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Tchernev, Velizar T.
; APPLICANT: Verne, Corine A.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Edinger, Schlowit
; APPLICANT: Peyman, John A.
; APPLICANT: Stone, David J.
; APPLICANT: Ellerman, Karen
; APPLICANT: Gangolli, Esha A.
; APPLICANT: Boldog, Terence L.
; APPLICANT: Colman, Steven D.
; APPLICANT: Eisen, Andrew J.
; APPLICANT: Liu, Xiaohong
; APPLICANT: Padigar, Muralidhar

```

```

; APPLICANT: Spaderna, Steven K.
; APPLICANT: Zehusen, Bryan D.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-216
; CURRENT APPLICATION NUMBER: US/10/231,913
; CURRENT FILING DATE: 2002-08-30
; PRIOR APPLICATION NUMBER: 60/251,660
; PRIOR FILING DATE: 2000-12-06
; PRIOR APPLICATION NUMBER: 60/255,029
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: 60/260,326
; PRIOR FILING DATE: 2001-01-08
; PRIOR APPLICATION NUMBER: 60/263,800
; PRIOR FILING DATE: 2001-01-24
; PRIOR APPLICATION NUMBER: 60/269,942
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/286,183
; PRIOR FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: 60/313,627
; PRIOR FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: 60/318,712
; PRIOR FILING DATE: 2001-09-12
; NUMBER OF SEQ ID NOS: 292
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 121
; LENGTH: 779
; TYPE: PRT
; ORGANISM: Mus musculus
; US-10-231-913-121

Query Match      38.1%; Score 48; DB 15; Length 779;
Best Local Similarity 52.9%; Pred. No. 1.2e+02;
Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY      5 SYDNFALVDYLVFERTK 21
DB      330 SYNHPALYYLLERLK 346

RESULT 5
US-09-823-187-26
; Sequence 26, Application US/09823187
; Publication No. US20030096952A1
; GENERAL INFORMATION:
; APPLICANT: Burgess, Catherine
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Liu, Xiaohong
; APPLICANT: Majumder, Kumud
; APPLICANT: Padigar, Muralidhar
; APPLICANT: Patturajan, Meera
; APPLICANT: Shinkete, Richard A
; APPLICANT: Spaderna, Steven K
; APPLICANT: Spytek, Kimberly
; APPLICANT: Taupier, Raymond J
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-745
; CURRENT APPLICATION NUMBER: US/09/823,187
; CURRENT FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/193,339
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 60/193,205
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 60/195,343
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: 60/195,088
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,005
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,792
; PRIOR FILING DATE: 2000-04-10
; PRIOR APPLICATION NUMBER: 60/196,556
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: 60/197,081

```

```

; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: 60/197,525
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/197,087
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 103
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 26
; LENGTH: 782
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-823-187-26

Query Match          38.1%; Score 48; DB 10; Length 782;
Best Local Similarity 52.9%; Pred. No. 1.2e+02;
Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 5 SYNFPALVDYLVFERTK 21
DB 330 SYNHPAIVYLLERLK 346

RESULT 6
US-09-815-915-2
; Sequence 2, Application US/09815915
; Patent No. US20020025931A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; TITLE OF INVENTION: 3714,16742, 23546, AND 13887 NOVEL
; FILE REFERENCE: 38155-20006.00
; CURRENT APPLICATION NUMBER: US/09/815,915
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 60/191,846
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 783
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-815-915-2

Query Match          38.1%; Score 48; DB 9; Length 783;
Best Local Similarity 52.9%; Pred. No. 1.2e+02;
Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 5 SYNFPALVDYLVFERTK 21
DB 330 SYNHPAIVYLLERLK 346

RESULT 7
US-09-823-187-90
; Sequence 90, Application US/09823187
; Publication No. US20030096952A1
; GENERAL INFORMATION:
; APPLICANT: Burgess, Catherine Y
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Liu, Xiaohong
; APPLICANT: Majumder, Kumud
; APPLICANT: Padigaru, Muralidhar
; APPLICANT: Paturajan, Meera
; APPLICANT: Shinkets, Richard A
; APPLICANT: Spaderna, Steven K
; APPLICANT: Spytek, Kimberly
; APPLICANT: Taupier, Raymond J
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-745
; CURRENT APPLICATION NUMBER: US/09/823,187
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/193,339
```

```

; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 60/193,205
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 60/195,343
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: 60/195,088
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,005
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,792
; PRIOR FILING DATE: 2000-04-10
; PRIOR APPLICATION NUMBER: 60/196,556
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: 60/197,081
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: 60/197,525
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 103
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 90
; LENGTH: 783
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-823-187-90

Query Match          38.1%; Score 48; DB 10; Length 783;
Best Local Similarity 52.9%; Pred. No. 1.2e+02;
Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 5 SYNFPALVDYLVFERTK 21
DB 330 SYNHPAIVYLLERLK 346

RESULT 8
US-10-393-316-2
; Sequence 2, Application US/10393316
; Publication No. US20030175786A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Meyers, Rachel
; TITLE OF INVENTION: 3714, 16742, 23546, AND 13887 NOVEL
; FILE REFERENCE: MP100-086P1RCNM
; CURRENT APPLICATION NUMBER: US/10/393,316
; CURRENT FILING DATE: 2003-03-20
; PRIOR APPLICATION NUMBER: US 09/815,915
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 60/191,846
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 783
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-393-316-2

Query Match          38.1%; Score 48; DB 14; Length 783;
Best Local Similarity 52.9%; Pred. No. 1.2e+02;
Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 5 SYNFPALVDYLVFERTK 21
DB 330 SYNHPAIVYLLERLK 346

RESULT 9
US-09-823-187-91
; Sequence 91, Application US/09823187
; Publication No. US20030096952A1
```

```
/ GENERAL INFORMATION:
/ APPLICANT: Burgess, Catherine
/ APPLICANT: Gusev, Vladimir Y
/ APPLICANT: Liu, Xiaohong
/ APPLICANT: Majumder, Kumud
/ APPLICANT: Padigaru, Muralidhar
/ APPLICANT: Patnrajan, Meera
/ APPLICANT: Shinkets, Richard A
/ APPLICANT: Spaderna, Steven K
/ APPLICANT: Spytek, Kimberly
/ APPLICANT: Taupier, Raymond J
/ TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
/ FILE REFERENCE: 15966-745
/ CURRENT APPLICATION NUMBER: US/09/823,187
/ CURRENT FILING DATE: 2001-03-29
/ PRIOR APPLICATION NUMBER: 60/193,339
/ PRIOR FILING DATE: 2000-03-30
/ PRIOR APPLICATION NUMBER: 60/193,205
/ PRIOR FILING DATE: 2000-03-30
/ PRIOR APPLICATION NUMBER: 60/195,343
/ PRIOR FILING DATE: 2000-04-05
/ PRIOR APPLICATION NUMBER: 60/195,088
/ PRIOR FILING DATE: 2000-04-06
/ PRIOR APPLICATION NUMBER: 60/195,005
/ PRIOR FILING DATE: 2000-04-06
/ PRIOR APPLICATION NUMBER: 60/195,792
/ PRIOR FILING DATE: 2000-04-10
/ PRIOR APPLICATION NUMBER: 60/196,556
/ PRIOR FILING DATE: 2000-04-11
/ PRIOR APPLICATION NUMBER: 60/197,081
/ PRIOR FILING DATE: 2000-04-13
/ PRIOR APPLICATION NUMBER: 60/197,525
/ PRIOR FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/197,087
/ PRIOR FILING DATE: 2000-04-14
/ NUMBER OF SEQ ID NOS: 103
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 91
/ LENGTH: 786
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ US-09-823-187-91

Query Match      38.1%; Score 48; DB 10; Length 786;
Best Local Similarity 52.9%; Pred. No. 1.2e+02;
Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy      5 SYNFAALVDYLVPERTK 21
Db      333 SYNFAALYLLERLK 349

RESULT 10
/ US-10-231-913-118
/ Sequence 118, Application US/10231913
/ Publication No. US20040005576A1
/ GENERAL INFORMATION:
/ APPLICANT: Guo, Xiaojia S.
/ APPLICANT: Li, Li
/ APPLICANT: Patnrajan, Meera
/ APPLICANT: Shinkets, Richard A.
/ APPLICANT: Casman, Stacie J.
/ APPLICANT: Malyankar, Uriel M.
/ APPLICANT: Tchernev, Vellizar T.
/ APPLICANT: Verneet, Corine A.
/ APPLICANT: Spytek, Kimberly A.
/ APPLICANT: Shenoy, Suresh G.
/ APPLICANT: Alsobrook II, John P.
/ APPLICANT: Edinger, Schlowit
/ APPLICANT: Peyman, John A.
/ APPLICANT: Stone, David J.
/ APPLICANT: Ellerman, Karen
/ APPLICANT: Gangolli, Eshna A.
```

```
/ APPLICANT: Boldog, Ference L.
/ APPLICANT: Colman, Steven D.
/ APPLICANT: Eissen, Andrew J.
/ APPLICANT: Liu, Xiaohong
/ APPLICANT: Padigaru, Muralidhar
/ APPLICANT: Spaderna, Steven K.
/ APPLICANT: Zerhusen, Bryan D.
/ TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
/ FILE REFERENCE: 21402-216
/ CURRENT APPLICATION NUMBER: US/10/231,913
/ CURRENT FILING DATE: 2002-08-30
/ PRIOR APPLICATION NUMBER: 60/251,660
/ PRIOR FILING DATE: 2000-12-06
/ PRIOR APPLICATION NUMBER: 60/255,029
/ PRIOR FILING DATE: 2000-12-12
/ PRIOR APPLICATION NUMBER: 60/260,326
/ PRIOR FILING DATE: 2001-01-08
/ PRIOR APPLICATION NUMBER: 60/263,800
/ PRIOR FILING DATE: 2001-01-24
/ PRIOR APPLICATION NUMBER: 60/269,942
/ PRIOR FILING DATE: 2001-02-20
/ PRIOR APPLICATION NUMBER: 60/286,183
/ PRIOR FILING DATE: 2001-04-24
/ PRIOR APPLICATION NUMBER: 60/313,627
/ PRIOR FILING DATE: 2001-08-20
/ PRIOR APPLICATION NUMBER: 60/318,712
/ PRIOR FILING DATE: 2001-09-12
/ NUMBER OF SEQ ID NOS: 292
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 118
/ LENGTH: 786
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ US-10-231-913-118

Query Match      38.1%; Score 48; DB 15; Length 786;
Best Local Similarity 52.9%; Pred. No. 1.2e+02;
Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy      5 SYNFAALVDYLVPERTK 21
Db      333 SYNFAALYLLERLK 349

RESULT 11
/ US-09-823-187-95
/ Sequence 95, Application US/09823187
/ Publication No. US20030096952A1
/ GENERAL INFORMATION:
/ APPLICANT: Burgess, Catherine
/ APPLICANT: Gusev, Vladimir Y
/ APPLICANT: Liu, Xiaohong
/ APPLICANT: Majumder, Kumud
/ APPLICANT: Padigaru, Muralidhar
/ APPLICANT: Patnrajan, Meera
/ APPLICANT: Shinkets, Richard A
/ APPLICANT: Spaderna, Steven K
/ APPLICANT: Spytek, Kimberly
/ APPLICANT: Taupier, Raymond J
/ TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
/ FILE REFERENCE: 15966-745
/ CURRENT APPLICATION NUMBER: US/09/823,187
/ CURRENT FILING DATE: 2001-03-29
/ PRIOR APPLICATION NUMBER: 60/193,339
/ PRIOR FILING DATE: 2000-03-30
/ PRIOR APPLICATION NUMBER: 60/193,205
/ PRIOR FILING DATE: 2000-03-30
/ PRIOR APPLICATION NUMBER: 60/195,343
/ PRIOR FILING DATE: 2000-04-05
/ PRIOR APPLICATION NUMBER: 60/195,088
/ PRIOR FILING DATE: 2000-04-06
/ PRIOR APPLICATION NUMBER: 60/195,005
/ PRIOR FILING DATE: 2000-04-06
```

PRIOR APPLICATION NUMBER: 60/195,792
PRIOR FILING DATE: 2000-04-10
PRIOR APPLICATION NUMBER: 60/196,556
PRIOR FILING DATE: 2000-04-11
PRIOR APPLICATION NUMBER: 60/197,081
PRIOR FILING DATE: 2000-04-13
PRIOR APPLICATION NUMBER: 60/197,525
PRIOR FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/197,087
PRIOR FILING DATE: 2000-04-14
NUMBER OF SEQ ID NOS: 103
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 95
LENGTH: 798
TYPE: PRT
ORGANISM: Gallus gallus
US-09-823-187-95

Query Match 38.1%; Score 48; DB 10; Length 798;
Best Local Similarity 52.9%; Pred. No. 1.3e+02;
Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 5 SYDNFALVDYLVPERK 21
DB 329 SYNHPAIIYLLERLK 345

RESULT 12
US-10-231-913-122
Sequence 122, Application US/10231913
Publication No. US2004005576A1
GENERAL INFORMATION:
APPLICANT: Guo, Xiaojia S.
APPLICANT: Li, Li
APPLICANT: Patturajan, Meera
APPLICANT: Shinkets, Richard A.
APPLICANT: Casman, Stacie J.
APPLICANT: Malyankar, Uriel M.
APPLICANT: Tchernev, Velizar T.
APPLICANT: Verneet, Corine A.
APPLICANT: Spytek, Kimberly A.
APPLICANT: Shenoy, Sureeh G.
APPLICANT: Alsobrook II, John P.
APPLICANT: Edinger, Schluomit
APPLICANT: Peyman, John A.
APPLICANT: Stone, David J.
APPLICANT: Ellerman, Karen
APPLICANT: Gangolli, Esha A.
APPLICANT: Boldog, Ference L.
APPLICANT: Colman, Steven D.
APPLICANT: Eisen, Andrew J.
APPLICANT: Liu, Xiaohong
APPLICANT: Padigaru, Muralidhara
APPLICANT: Spaderna, Steven K.
APPLICANT: Zerhusen, Bryan D.
TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-216
CURRENT APPLICATION NUMBER: US/10/231,913
CURRENT FILING DATE: 2002-08-30
PRIOR APPLICATION NUMBER: 60/251,660
PRIOR FILING DATE: 2000-12-06
PRIOR APPLICATION NUMBER: 60/255,029
PRIOR FILING DATE: 2000-12-12
PRIOR APPLICATION NUMBER: 60/260,326
PRIOR FILING DATE: 2001-01-08
PRIOR APPLICATION NUMBER: 60/263,800
PRIOR FILING DATE: 2001-01-24
PRIOR APPLICATION NUMBER: 60/269,942
PRIOR FILING DATE: 2001-02-20
PRIOR APPLICATION NUMBER: 60/266,183
PRIOR FILING DATE: 2001-04-24
PRIOR APPLICATION NUMBER: 60/313,627
PRIOR FILING DATE: 2001-08-20

PRIOR APPLICATION NUMBER: 60/318,712
PRIOR FILING DATE: 2001-09-12
NUMBER OF SEQ ID NOS: 292
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 122
LENGTH: 798
TYPE: PRT
ORGANISM: Gallus gallus
US-10-231-913-122

Query Match 38.1%; Score 48; DB 15; Length 798;
Best Local Similarity 52.9%; Pred. No. 1.3e+02;
Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 5 SYDNFALVDYLVPERK 21
DB 329 SYNHPAIIYLLERLK 345

RESULT 13
US-10-425-114-54182
Sequence 54182, Application US/10425114
Publication No. US20040034888A1
GENERAL INFORMATION:
APPLICANT: Liu, Jingdong
APPLICANT: Zhou, Yihua
APPLICANT: Kovacic, David K.
APPLICANT: Screen, Steven E
APPLICANT: Tabaka, Jack E
APPLICANT: Cao, Yongwei

TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 38-21(53313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 54182
LENGTH: 826
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Clone ID: LIB4119-026-D11_F11.pep
US-10-425-114-54182

Query Match 38.1%; Score 48; DB 15; Length 826;
Best Local Similarity 52.9%; Pred. No. 1.3e+02;
Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 5 SYDNFALVDYLVPERK 21
DB 373 SYNHPAIIYLLERLK 389

RESULT 14
US-10-437-963-201611
Sequence 201611, Application US/10437963
Publication No. US20040123343A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovacic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 201611

LENGTH: 1193
TYPE: PRT
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_9696C.1.pep
US-10-437-963-201611

Query Match 38.1%; Score 48; DB 16; Length 1193;
Best Local Similarity 42.9%; Pred. No. 2e+02;
Matches 9; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY 2 ISVSNDPALVDYLVFERTKS 22
DB 1014 IDVGDDISGIDFVFERPES 1034

RESULT 15
US-09-764-868-834
Sequence 834, Application US/09764868
Patent No. US20020168711A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PT232
CURRENT APPLICATION NUMBER: US/09/764,868
CURRENT FILING DATE: 2001-01-17
Prior application data removed - refer to PALM or file wrapper
NUMBER OF SEQ ID NOS: 1510
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 834
LENGTH: 511
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (236)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (244)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (278)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (330)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (406)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (421)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (460)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (470)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-868-834

Query Match 37.3%; Score 47; DB 9; Length 511;
Best Local Similarity 50.0%; Pred. No. 1.1e+02;
Matches 9; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 5 SYDNFALVDYLVFERTKS 22
DB 122 SYNHPALYFLVLRKLS 139

RESULT 16
US-09-764-875-636
Sequence 636, Application US/09764875
Publication No. US20040018965A1

GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: P202
CURRENT APPLICATION NUMBER: US/09/764,875
CURRENT FILING DATE: 2001-01-17
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 1249
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 636
LENGTH: 522
TYPE: PRT
ORGANISM: Homo sapiens
US-09-764-875-636

Query Match 37.3%; Score 47; DB 11; Length 522;
Best Local Similarity 50.0%; Pred. No. 1.1e+02;
Matches 9; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 5 SYDNFALVDYLVFERTKS 22
DB 122 SYNHPALYFLVLRKLS 139

RESULT 17
US-10-425-114-39113
Sequence 39113, Application US/10425114
Publication No. US20040034888A1
GENERAL INFORMATION:
APPLICANT: Liu, Jingdong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E.
APPLICANT: Tabaska, Jack E.
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 38-21(5313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 39113
LENGTH: 589
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Clone ID: LIB3063-047-F10_FLI.pep
US-10-425-114-39113

Query Match 37.3%; Score 47; DB 15; Length 589;
Best Local Similarity 50.0%; Pred. No. 1.3e+02;
Matches 9; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 5 SYDNFALVDYLVFERTKS 22
DB 93 SYNHPALYFLVLRKLS 110

RESULT 18
US-10-408-765A-1811
Sequence 1811, Application US/10408765A
Publication No. US20040101874A1
GENERAL INFORMATION:
APPLICANT: Ghosh, Soumitra S.
APPLICANT: Fahy, Bo D.
APPLICANT: Zhang, Bing
APPLICANT: Gibson, Bradford W.
APPLICANT: Taylor, Steven W.
APPLICANT: Glenn, Gary M.
APPLICANT: Warnock, Dale E.
TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
FILE REFERENCE: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
US-10-408-765A-1811

;; CURRENT APPLICATION NUMBER: US/10/408,765A
;; CURRENT FILING DATE: 2003-04-04
;; NUMBER OF SEQ ID NOS: 3077
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 1811
;; LENGTH: 896
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-408-765A-1811

Query Match 37.3%; Score 47; DB 16; Length 896;
Best Local Similarity 50.0%; Pred. No. 2.1e+02;
Matches 9; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Qy 5 SYDNFALVDYLVERTKS 22
Db 292 SYNHPAIFYLVLRLKS 309

RESULT 19
US-10-258-106-7
; Sequence 7, Application US/10258106
; Publication No. US20040018185A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.; YUE, Henry
; APPLICANT: GANDHI, Ameena R.; TRIBOULEY, Catherine M.
; APPLICANT: KEARNEY, Liam; GRIFFIN, Jennifer A.
; APPLICANT: NGUYEN, Daniel B.; BANDMAN, Olga
; APPLICANT: LU, Dung Alpha M.; LAU, Preeti G.
; APPLICANT: BURFORD, Neil; KHAN, Parrah A.
; APPLICANT: WALIA, Nandinder K.; YAO, Montique G.
; APPLICANT: PATTERSON, Chandra; BURRILL, John D.
; APPLICANT: MARCUS, Gregory A.; ZINGLER, Kurt A.
; APPLICANT: RECIPON, Shirley A.; LU, Yan
; APPLICANT: POLICKY, Jennifer L.; THORNTON, Michael B.
; APPLICANT: TANG, Y. Tom; HAFALIA, April J.A.
; APPLICANT: ELLIOTT, Vicki S.; BAUGHN, Marian R.
; APPLICANT: WALSH, Roderick T.; RANKUMAR, Jayalaxmi
; APPLICANT: BOROWSKY, Mark L.; AU-YOUNG, Janice K.
; APPLICANT: HILLMAN, Jennifer L.; GURURAJAN, Rajagopal
; TITLE OF INVENTION: HUMAN KINASES
; FILE REFERENCE: PI-0076 USN
; CURRENT APPLICATION NUMBER: US/10/258,106
; CURRENT FILING DATE: 2002-10-18
; PRIOR APPLICATION NUMBER: PCT/US01/12992
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: US 60/199,021
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/200,226
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/202,339
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/203,505
; PRIOR FILING DATE: 2000-05-11
; PRIOR APPLICATION NUMBER: US 60/205,654
; PRIOR FILING DATE: 2000-05-18
; PRIOR APPLICATION NUMBER: US 60/207,739
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 60/208,795
; PRIOR FILING DATE: 2000-06-01
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PERL Program
; SEQ ID NO 7
; LENGTH: 923
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20040018185A1 2589355CD1
US-10-258-106-7

Query Match 37.3%; Score 47; DB 15; Length 923;
Best Local Similarity 50.0%; Pred. No. 2.1e+02;

Matches 9; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
Qy 5 SYDNFALVDYLVERTKS 22
Db 319 SYNHPAIFYLVLRLKS 336

RESULT 20
US-10-437-963-149049
; Sequence 149049, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovacic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 149049
; LENGTH: 1673
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1673)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_4941C.1.pep
US-10-437-963-149049

Query Match 37.3%; Score 47; DB 16; Length 1673;
Best Local Similarity 40.7%; Pred. No. 4.3e+02;
Matches 11; Conservative 6; Mismatches 6; Indels 4; Gaps 2;

Qy 1 KISVSY-DNFALV---DYLVERTKSD 23
Db 875 KVMEYLDNFVVVFIDILVSRTEE 901

RESULT 21
US-08-754-311B-2
; Sequence 2, Application US/08754311B
; Publication No. US20020004221A1
; GENERAL INFORMATION:
; APPLICANT: Bonini, Nancy M.
; APPLICANT: Leiserson, William M.
; APPLICANT: Benzer, Seymour
; TITLE OF INVENTION: PROGRAMMED CELL DEATH ANTAGONIST
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/754,311B
; FILING DATE: 21-NOV-1996

Best Local Similarity 50.0%; Pred. No. 1.1e+02;
Matches 11; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

Db 375 VGLDNFDLVKYLVSQVWLSEED 396

RESULT 25

US-10-282-122A-56616
 ; Sequence 56616, Application US/10282122A
 ; Publication No. US20040029129A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Liangsu
 ; APPLICANT: Zamudio, Carlos
 ; APPLICANT: Malone, Cheryl
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Kari
 ; APPLICANT: Zyskind, Judith
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John
 ; APPLICANT: Carr, Grant
 ; APPLICANT: Yamamoto, Robert
 ; APPLICANT: Forsyth, R.
 ; APPLICANT: Xu, H.
 ; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 ; FILE REFERENCE: EPIGRA.034A
 ; CURRENT FILING DATE: 2003-02-20
 ; PRIOR APPLICATION NUMBER: 60/191,078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206,848
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR APPLICATION NUMBER: 60/207,727
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: 60/230,335
 ; PRIOR FILING DATE: 2000-09-06
 ; PRIOR APPLICATION NUMBER: 60/230,347
 ; PRIOR FILING DATE: 2000-09-09
 ; PRIOR APPLICATION NUMBER: 60/242,578
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/253,625
 ; PRIOR FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: 60/257,931
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: 60/267,636
 ; PRIOR FILING DATE: 2001-02-09
 ; PRIOR APPLICATION NUMBER: 60/269,308
 ; PRIOR FILING DATE: 2001-02-16
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 78614
 ; SOFTWARE: Patentin version 3.1
 ; SEQ ID NO 56616
 ; LENGTH: 548
 ; TYPE: PRT
 ; ORGANISM: Escherichia coli
 US-10-282-122A-56616

Query Match 36.5%; Score 46; DB 15; Length 548;
 Best Local Similarity 50.0%; Pred. NO. 1.7e+02;
 Matches 11; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 4 VSYDNFALVDYLVFERTKSDTD 25
 Db 375 VGLDNFDLVKYLVSQVWLSEED 396

Search completed: May 9, 2005, 12:31:14
 Job time : 133 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2005, 12:10:22 ; Search time 170 Seconds

(without alignments)
75.306 Million cell updates/sec

Title: US-10-694-847-3

Perfect score: 126

Sequence: 1 KISVSVDNPAUVDYLVPERKSDPTD 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	98.5	78.2	160	1 TRIA TRIA	Q27049 triatoma pa
2	50	39.7	223	2 Q8MOT9	Q7PVE2 anopheles g
3	50	39.7	223	2 Q7PVS2	Q64523 arabidopsis
4	49	38.9	140	2 Q64523	Q92VX3 rhizobium m
5	49	38.9	317	2 Q92VX3	Q92VX3 rhizobium m
6	49	38.9	346	2 Q92VX3	Q92VX3 rhizobium m
7	49	38.9	496	2 Q92VX3	Q92VX3 rhizobium m
8	49	38.9	512	1 MQO_BRAA	Q99XN4 bradyrhizob
9	49	38.9	914	2 Q9L3F1	Q913F1 rumicoccu
10	48	38.1	463	2 Q8C3I3	Q8C3I3 mus musculu
11	48	38.1	764	2 Q6ZNL8	Q6ZNL8 mus musculu
12	48	38.1	779	1 SNIL MOUSE	Q60670 mus musculu
13	48	38.1	783	2 Q86VT2	Q86VT2 mus musculu
14	48	38.1	786	1 SNIL HUMAN	PS7059 homo sapien
15	48	38.1	788	2 Q91A88	Q91A88 homo sapien
16	48	38.1	1193	2 Q91A88	Q91A88 homo sapien
17	47	37.3	232	2 Q6FS04	Q6FS04 gryza sativ
18	47	37.3	321	2 Q97AL5	Q97AL5 candida gla
19	47	37.3	330	2 Q05583	Q05583 thermoplas
20	47	37.3	346	2 Q8NCV7	Q8NCV7 homo sapien
21	47	37.3	505	2 Q9BZT1	Q9BZT1 meloidogyne
22	47	37.3	518	1 MALF_PHOLL	Q7N984 photorhabdu
23	47	37.3	527	2 Q6D2L3	Q6D2L3 erwinia car
24	47	37.3	634	2 Q88S13	Q88S13 lactobacill
25	47	37.3	711	2 Q6AZR2	Q6AZR2 homo sapien
26	47	37.3	926	2 Q9H0K1	Q9H0K1 homo sapien
27	47	37.3	931	2 Q8CFH6	Q8CFH6 mus musculu
28	47	37.3	950	2 Q76N03	Q76N03 homo sapien
29	47	37.3	986	2 Q7R9N7	Q7R9N7 plasmodium
30	47	37.3	1249	2 Q8REU9	Q8REU9 fusobacteri
31	47	37.3	2705	2 Q9W6V6	Q9W6V6 gallus gall

32	46.5	36.9	327	2 Q87W14	Q87W14 pseudomonas
33	46.5	36.9	552	2 Q91CM0	Q91CM0 antheraea p
34	46.5	36.9	892	2 Q91ZG4	Q91ZG4 cache valle
35	46	36.5	173	2 Q7WX23	Q7WX23 alcaligenes
36	46	36.5	228	2 Q7QYH9	Q7QYH9 giardia lam
37	46	36.5	335	2 Q6CG52	Q6CG52 yarrowia li
38	46	36.5	405	2 Q92F81	Q92F81 listeria in
39	46	36.5	408	2 Q8YAE4	Q8YAE4 listeria mo
40	46	36.5	408	2 Q724M7	Q724M7 listeria mo
41	46	36.5	426	2 Q6D901	Q6D901 erwinia car
42	46	36.5	520	2 Q80ZG0	Q80ZG0 ratuus norv
43	46	36.5	526	2 Q6KHD8	Q6KHD8 mycoplasma
44	46	36.5	548	1 MQO_ECOS7	Q8X45 escherichia
45	46	36.5	548	1 MQO_ECOS7	Q8X45 escherichia
46	46	36.5	548	1 MQO_ECOS7	Q8X45 escherichia
47	46	36.5	552	2 Q8A5E2	Q8A5E2 bacteroides
48	46	36.5	560	2 Q80X61	Q80X61 mus musculu
49	46	36.5	562	1 MQO_XYLF4	Q9PE6 xylella fas
50	46	36.5	562	1 MQO_XYLF4	Q9PE6 xylella fas
51	46	36.5	562	1 HGP2_HAEIN	P44809 haemophilus
52	45.5	36.1	158	1 V235_MYCPN	P75456 mycoplasma
53	45.5	36.1	674	2 Q9X2G0	Q9X2G0 thermotoga
54	45.5	36.1	897	2 Q997B5	Q997B5 cache valle
55	45.5	36.1	897	2 Q997B6	Q997B6 cache valle
56	45.5	36.1	897	2 Q997B7	Q997B7 cache valle
57	45.5	36.1	897	2 Q997B8	Q997B8 cache valle
58	45.5	36.1	897	2 Q997B9	Q997B9 cache valle
59	45.5	36.1	897	2 Q997B9	Q997B9 cache valle
60	45.5	36.1	1434	2 Q9J6L5	Q9J6L5 cache valle
61	45.5	36.1	1434	2 Q9J6L7	Q9J6L7 cache valle
62	45.5	36.1	1434	2 Q9PXS4	Q9PXS4 cache valle
63	45	35.7	131	2 Q7T5J3	Q7T5J3 cryptophleb
64	45	35.7	155	2 Q83EB3	Q83EB3 coxiella bu
65	45	35.7	357	2 Q6H0Y7	Q6H0Y7 sulfolobus
66	45	35.7	363	2 Q9A508	Q9A508 caulobacter
67	45	35.7	459	2 Q8DSH1	Q8DSH1 streptococc
68	45	35.7	530	2 Q953M3	Q953M3 echinococcu
69	45	35.7	776	1 SNIL_RAT	Q91407 rattus norv
70	45	35.7	1116	1 MKH1_SCHPO	Q10407 schistosach
71	45	35.7	1267	2 Q81LD1	Q81LD1 plasmodium
72	45	35.7	1895	2 Q7R1I4	Q7R1I4 plasmodium
73	44.5	35.3	122	2 Q630U8	Q630U8 bacillus ce
74	44.5	35.3	122	2 Q8DK00	Q8DK00 bacillus an
75	44.5	35.3	171	2 Q8E345	Q8E345 streptococc
76	44.5	35.3	171	2 Q8E345	Q8E345 streptococc
77	44.5	35.3	458	2 Q815U7	Q815U7 plasmodium
78	44.5	35.3	477	2 Q834L6	Q834L6 enterococcu
79	44.5	35.3	515	1 BRES_YEAST	PS3741 salmonella
80	44.5	35.3	1249	2 Q820Z3	Q820Z3 salmonella
81	44.5	35.3	1744	2 Q86B80	Q86B80 drosoephila
82	44.5	35.3	4092	1 DYHC_YEAST	P36022 saccharomyc
83	44	34.9	112	2 Q894C3	Q894C3 clostridium
84	44	34.9	139	2 Q85236	Q85236 lactobacill
85	44	34.9	222	2 Q8GJC7	Q8GJC7 campylobact
86	44	34.9	224	1 PLS2_HUMAN	Q9N977 homo sapien
87	44	34.9	246	2 Q9LJ66	Q9LJ66 arabisdopsi
88	44	34.9	278	2 Q8YXU9	Q8YXU9 arabisdopsi
89	44	34.9	286	2 Q6CPB7	Q6CPB7 kluyveromyc
90	44	34.9	299	2 Q6NSW9	Q6NSW9 homo sapien
91	44	34.9	317	2 Q724L7	Q724L7 homo sapien
92	44	34.9	319	2 Q7VON6	Q7VON6 prochloroco
93	44	34.9	324	2 Q9NFA4	Q9NFA4 plasmodium
94	44	34.9	324	2 Q9NFA4	Q9NFA4 plasmodium
95	44	34.9	331	2 Q9KD62	Q9KD62 bacillus ha
96	44	34.9	331	2 Q9KD62	Q9KD62 bacillus ha
97	44	34.9	434	1 Q8IKP4	P21026 vaccinia vi
98	44	34.9	434	1 VG05_VACCC	P32995 vaccinia vi
99	44	34.9	434	1 VG05_VACCC	P32995 vaccinia vi
100	44	34.9	434	2 Q8JL65	Q8JL65 ectromelia

ALIGNMENTS

RESULT 1
 ID TRIA TRIA STANDARD; PRT; 160 AA.
 AC Q27049; Q27046; Q27047; Q27048;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Triabin precursor.
 OS Triatoma pallidipennis (Triatomine bug).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Paraneoptera; Hemiptera; Euhemiptera; Heteroptera;
 OC Panheteroptera; Cimicomorpha; Reduviidae; Triatominae; Triatoma.
 NCBI_TaxID=30077;
 RN [1]
 RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
 RC TRIA TRIA TRIA gland;
 RX MEDLINE=96081922; PubMed=7499380; DOI=10.1074/jbc.270.48.28629;
 RA Noeske-Jungblut C., Heendler B., Donner P., Alagon A., Possant L.D.,
 RA Schleuning W.D.;
 RT "Triabin, a highly potent exosite inhibitor of thrombin.";
 RL J. Biol. Chem. 270:28629-28634(1995).
 RN [2]
 RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF COMPLEX WITH THROMBIN.
 RX MEDLINE=98004486; PubMed=9342325; DOI=10.1073/pnas.94.22.11845;
 RA Fuentes-Prior P., Noeske-Jungblut C., Donner P., Schleuning W.D.,
 RA Huber R., Bode W.;
 RT "Structure of the thrombin complex with triabin, a lipocalin-like
 RT exosite-binding inhibitor derived from a triatomine bug.";
 RL Proc. Natl. Acad. Sci. U.S.A. 94:11845-11850(1997).
 CC -1- FUNCTION: Thrombin inhibitor. Forms a noncovalent complex with
 CC thrombin at a molar ratio of 1:1, inhibits thrombin-induced
 CC platelet aggregation, and prolongs thrombin clotting time and
 CC activated partial thromboplastin time. It only mutually
 CC suppresses the amidolytic activity of thrombin.
 CC -1- SUBCELLULAR LOCATION: Secreted.
 CC -1- MISCELLANEOUS: The sequence shown is that of clone TR5.
 CC -1- SIMILARITY: Belongs to the lipocalin family. Triabin subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X80246; CA55540.1; -;
 DR EMBL; X80247; CA55541.1; -;
 DR EMBL; X80248; CA55542.1; -;
 DR EMBL; X80249; CA55543.1; -;
 DR PDB; IAVG; X-ray; I-19-160.
 DR InterPro: IPR011038; Calycin.
 DR InterPro: IPR005657; Triabin.
 DR Pfam; PF03973; Triabin; 1.
 DR ProDom; PD403487; Lipocalin like; 1.
 DR 3D-structure; Blood coagulation; Direct protein sequencing; Lipocalin;
 KM Serine protease inhibitor; Signal.
 FT STGNAL 1 18
 FT CHAIN 19 160
 FT DISULFID 24 128
 FT DISULFID 57 160
 FT DISULFID 87 102
 FT VARIANT 26 26
 FT VARIANT 50 50
 FT VARIANT 90 95
 FT VARIANT 95 95
 FT VARIANT 104 104
 FT VARIANT 132 132
 FT VARIANT 145 145
 FT VARIANT 157 157
 FT TURN 20 21
 FT HELIX 24 26
 I -> L (in clones TR12, TR28 and TR45).
 G -> D (in clones TR2 and TR45).
 VDNKNG -> ADKND (in clone TR45).
 S -> D (in clones TR12 and TR28).
 S -> G (in clones TR12, TR28 and TR45).
 T -> I (in clones TR12 and TR28).
 L -> F (in clones TR12, TR28 and TR45).
 K -> N (in clones TR12, TR28 and TR45).

FT HELIX 35 38
 FT STRAND 42 46
 FT STRAND 55 61
 FT STRAND 69 74
 FT TURN 77 78
 FT TURN 81 82
 FT STRAND 84 89
 FT TURN 94 96
 FT STRAND 98 104
 FT TURN 105 106
 FT TURN 110 119
 FT STRAND 121 122
 FT STRAND 124 133
 FT STRAND 140 146
 FT STRAND 140 146
 FT HELIX 156 160
 SQ SEQUENCE 160 AA; 17860 MW; 7DB62819F7C4406A CRC64;
 Query Match 78.2%; Score 98.5; DB 1; Length 160;
 Best Local Similarity 62.2%; Pred. No. 5.7e-07;
 Matches 23; Conservative 0; Mismatches 1; Indels 13; Gaps 1;
 QY 2 ISVSYNFALV-----DYLVFERTSDTD 25
 DB 117 ISVSYNFALVCRSITFTSQPKEDYLVFERTSDTD 153
 RESULT 2
 ID Q8WQJ9 PRELIMINARY; PRT; 223 AA.
 AC Q8WQJ9;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Glutathione S-transferase E3.
 GN Name=GSTe3;
 OS Anopheles gambiae (African malaria mosquito).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.
 NCBI_TaxID=7165;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ZAN/U;
 RX MEDLINE=22755733; PubMed=12718742; DOI=10.1042/BJ20030169;
 RA Ostell F., Rosalier L.C., Vontas J., Ranson H., Hwangway J.;
 RT "Heterologous expression of four glutathione transferase genes
 RT genetically linked to a major insecticide-resistance locus from the
 RT malaria vector Anopheles gambiae.";
 RL Biochem. J. 373:957-963(2003).
 CC -1- SIMILARITY: Belongs to the GST superfamily.
 DR EMBL; AY070234; AAL58538.1; -;
 DR HSSP; Q93113; 1PN9.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR Pfam; PF00043; GST_C; 1.
 DR Pfam; PF02798; GST_N; 1.
 DR TRANSFERASE.
 KM SEQUENCE 223 AA; 25156 MW; 79F262BBE682BC89 CRC64;
 SQ
 Query Match 39.7%; Score 50; DB 2; Length 223;
 Best Local Similarity 47.4%; Pred. No. 23;
 Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;
 QY 6 YDNFALVDYLVFERTSDT 24
 DB 66 YDSHAIINLVQKAKDDT 84
 RESULT 3
 ID Q7PVS2 PRELIMINARY; PRT; 223 AA.
 AC Q7PVS2;
 DT 01-MAR-2004 (TrEMBLrel. 26, Created)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE ENSANGP0000016648.
GN Name=ENSANG00000014159;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoides; Anopheles.
RN NCBI_TaxID=180454;
RP [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the GST superfamily.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAAB01008984; EAA14730.2; -.
DR HSP; Q9113; 1PN9.
DR InterPro; IPR004046; GST_Cterm.
DR InterPro; IPR010987; GST_C like.
DR InterPro; IPR004045; GST_Nterm.
DR Pfam; PF00043; GST_C; 1.
DR Pfam; PF02798; GST_N; 1.
SQ SEQUENCE 223 AA; 25156 MW; 79F262EBE682BC98 CRC64;

Query Match 39.7%; Score 50; DB 2; Length 223;
Best Local Similarity 47.4%; Pred. No. 23;
Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 6 YDNFALVDYLVFERTKSDT 24
Db 66 YDSHAIINYLQKAKADDT 84

RESULT 4
064523 PRELIMINARY; PRT; 140 AA.
ID 064523;
AC 064523;
DT 01-AUG-1998 (TEMBLrel. 07, Created)
DT 01-AUG-1998 (TEMBLrel. 07, Last sequence update)
DT 01-JUN-2003 (TEMBLrel. 24, Last annotation update)
DE YUP8H12R.10.
GN Name=YUP8H12R.10;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC Eurosid II; Brassicales; Brassicaceae; Arabidopsis.
RN NCBI_TaxID=3702;
RP [1]
RP SEQUENCE FROM N.A.
RA Theologis A., Vysotskaia V.S., Osborne B.I., Schwartz J.R.,
RA Federspiel N.A., Kwan A., Toriumi M., Yu G., Oji O., Araujo R.,
RA Chung E., Dewar K., Dietrich F., Ecker J.R., Marzilli A., Oefner P.,
RA Davis R.W.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Theologis A.
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC002986; AAC17059.1; -.
DR PIR; T01027; T01027.
SQ SEQUENCE 140 AA; 16133 MW; 02BC26A9E1D11E2 CRC64;

Query Match 38.9%; Score 49; DB 2; Length 140;
Best Local Similarity 43.8%; Pred. No. 20;
Matches 7; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

RESULT 5
Q92VX3

ID Q92VX3 PRELIMINARY; PRT; 317 AA.
AC Q92VX3;
DT 01-DEC-2001 (TEMBLrel. 19, Created)
DT 01-DEC-2001 (TEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TEMBLrel. 24, Last annotation update)
DE Putative polysaccharide export-associated protein.
GN ORFNames=Smb20831;
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Plasmid pSymb.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
RN NCBI_TaxID=382;
RP [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396508; PubMed=11481431; DOI=10.1073/pnas.161294698;
RA Finan T.M., Weidner S., Wong K., Bumester J., Chain P.,
RA Vorhoeiter F.J., Hernandez-Lucas I., Becker A., Gouzy J.,
RA Golding B., Puehler A.;
RL "The complete sequence of the 1,683-kb pSymb megaplasmid from the N2-
RT fixing endosymbiont Sinorhizobium meliloti."
RL Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894(2001).
DR EMBL; AL591985; CAC48974.1; -.
DR PIR; F95913; F95913.
KM Complete proteome; Plasmid.
SQ SEQUENCE 317 AA; 35136 MW; F443FEF661D0816 CRC64;

QY 7 DNFALVDYLVFERTKSDT 25
Db 13 DSYIMDYIUSERMADAD 31

RESULT 6
09R9MO PRELIMINARY; PRT; 346 AA.
ID 09R9MO;
AC 09R9MO;
DT 01-MAY-2000 (TEMBLrel. 13, Created)
DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
DE Putative polysaccharide export protein.
GN Name-rkp;
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
RN NCBI_TaxID=382;
RP [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Rm41;
RX MEDLINE=21619047; PubMed=11768534;
RA Kils E., Kereszt A., Barta F., Stephens S., Reus B., Kondorosi A.,
RA Putnoky P.;
RL "The rkp-3 gene region of Sinorhizobium meliloti Rm41 contains strain-
RT specific genes that determine K antigen structure."
RL Mol. Plant Microbe Interact. 14:1395-1403(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Rm41;
RX PubMed=1496788; DOI=10.1128/JB.186.6.1591-1597.2004;
RA Putnoky P., Deak V., Bekasi K., Palvolgyi A., Maasz A., Palagyi Z.,
RA Hofmann G., Kerepesi I.;
RL "H protein of bacteriophage 16-3 and RkpM protein of Sinorhizobium
RT meliloti 41 are involved in phage adsorption."
RL J. Bacteriol. 186:1591-1597(2004).
DR EMBL; AJ245666; CAB62156.1; -.
SQ SEQUENCE 346 AA; 38442 MW; 58FA1D951F672B30 CRC64;

Query Match 38.9%; Score 49; DB 2; Length 346;
Best Local Similarity 42.1%; Pred. No. 51;
Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 7 DNFALVDYLVFERTSDTD 25
 DB 43 DSVLMDYILISERMAAD 61

RESULT 7

MOO_PROMM STANDARD; PRT; 496 AA.
 AC 07V8S6;
 DT 29-MAR-2004 (Rel. 43, Created)
 DT 29-MAR-2004 (Rel. 43, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Probable malate:quinone oxidoreductase (EC 1.1.99.16) (Malate
 dehydrogenase [acceptor]) (MOO).
 GN Name=mgo; OrderedLocustNames=PMT0255;
 OS Prochlorococcus marinus (strain MIT 9313).
 OC Bacteria; Cyanobacteria; Prochlorales; Prochlorococaceae;
 OC Prochlorococcus.
 NCBI_TaxID=74547;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=2285698; PubMed=12917642; DOI=10.1038/nature01947;
 RA Rocap G., Larimer F.W., Lamerdin J.B., Malfatti S., Chain P.,
 RA Algren N.A., Ariello A., Coleman M., Hauser L., Hess W.R.,
 RA Johnson Z.I., Land M.L., Lindell D., Post A.F., Regala W., Shah M.,
 RA Shaw S.L., Steglich C., Sullivan M.B., Ting C.S., Tolonen A.,
 RA Webb E.A., Zinser E.R., Chisholm S.W.;
 RT "genome divergence in two Prochlorococcus ecotypes reflects oceanic
 niche differentiation."
 RL Nature 424:1042-1047(2003).
 CC -1- CATALYTIC ACTIVITY: (S)-malate + acceptor = oxaloacetate + reduced
 acceptor.
 CC -1- COFACTOR: FAD (By similarity).
 CC -1- PATHWAY: Tricarboxylic acid cycle.
 CC -1- SIMILARITY: Belongs to the MOO family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: BX572095; CAB20430.1; -.
 DR HAMAP: MF_00212; -1.
 DR InterPro: IPR006231; mal_quin_oxido.
 DR Pfam: PF06039; Mgo; 1.
 DR TIGRFAMs: TIGR01320; mal_quin_oxido; 1.
 KM Complete proteome; FAD; Flavoprotein; Oxidoreductase;
 KM Tricarboxylic acid cycle.
 SQ SEQUENCE 496 AA; 53899 MW; D32F3D17FDC08E9 CRC64;

Query Match 38.9%; Score 49; DB 1; Length 496;
 Best Local Similarity 41.7%; Pred. No. 74;
 Matches 10; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 2 ISVSYDNFALVDYLVFERTSDTD 25
 DB 351 LEVGFKNFDLVLYLSELQOSEKD 374

RESULT 8

MOO_BRAJA STANDARD; PRT; 512 AA.
 AC 089XM4;
 DT 29-MAR-2004 (Rel. 43, Created)
 DT 29-MAR-2004 (Rel. 43, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Probable malate:quinone oxidoreductase (EC 1.1.99.16) (Malate
 dehydrogenase [acceptor]) (MOO).
 GN Name=mgo; OrderedLocustNames=b110284;

OS Bradyrhizobium japonicum.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Bradyrhizobiaceae; Bradyrhizobium.
 NCBI_TaxID=375;

RESULT 9

RP SEQUENCE FROM N.A.
 RC STRAIN=USDA 110;
 RX MEDLINE=22484998; PubMed=12597275;
 RA Kaneke T., Nakamura Y., Sato S., Minamisawa K., Uchiyama T.,
 RA Sasamoto S., Matsumoto A., Idesawa K., Iriyuchi M., Kawashima K.,
 RA Kohara M., Matsumoto M., Shimo S., Tsuruka H., Wada T., Yamada M.,
 RA Tabata S.;
 RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
 Bradyrhizobium japonicum USDA110."
 RL DNA Res. 9:189-197(2002).
 CC -1- CATALYTIC ACTIVITY: (S)-malate + acceptor = oxaloacetate + reduced
 acceptor.
 CC -1- COFACTOR: FAD (By similarity).
 CC -1- PATHWAY: Tricarboxylic acid cycle.
 CC -1- SIMILARITY: Belongs to the MOO family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AP005935; BAC45549.1; -.
 DR HAMAP: MF_00212; -1.
 DR InterPro: IPR006231; mal_quin_oxido.
 DR Pfam: PF06039; Mgo; 1.
 DR TIGRFAMs: TIGR01320; mal_quin_oxido; 1.
 KM Complete proteome; FAD; Flavoprotein; Oxidoreductase;
 KM Tricarboxylic acid cycle.
 SQ SEQUENCE 512 AA; 55455 MW; 5B0FF5455898F765 CRC64;

Query Match 38.9%; Score 49; DB 1; Length 512;
 Best Local Similarity 43.5%; Pred. No. 76;
 Matches 10; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

QY 2 ISVSYDNFALVDYLVFERTSDTD 24
 DB 350 LSVARDNFDLTXLVGVQLQSES 372

RA Gomez-Rodriguez A.;
 RL Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ276653; CAB93674.2; -
 SQ SEQUENCE 914 AA; 107410 MW; 68FBBF7963191690 CRC64;
 Query Match 38.9%; Score 49; DB 2; Length 914;
 Best Local Similarity 43.5%; Pred. No. 1.4e+02;
 Matches 10; Conservative 5; Mismatches 8; Indels 0; Gaps 0;
 QY 3 SVSYNPAVDLYVFERKSDTD 25
 Db 831 SLCHGNFALIDFLISYRKIVGTD 853
 RESULT 10
 Q8C313 PRELIMINARY; PRT; 463 AA.
 AC Q8C313;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Mus musculus 0 day neonate lung cDNA, RIKEN full-length enriched
 DE library, clone:E030044014 product:SNF1-like kinase, full insert
 DE sequence. (Fragment).
 GN Name=Snf1k;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 ON NCB1_TaxID=10090;
 RX SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Lung;
 RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
 RA Carninci P., Hayashizaki Y.;
 RT "High-efficiency full-length cDNA cloning."; Meth. Enzymol. 303:19-44(1999).
 RL Nature 409:685-690(2001).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Lung;
 RA The FANTOM Consortium;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs."; Nature 420:563-573(2002).
 RL [4]
 RN SEQUENCE FROM N.A.
 RP STRAIN=C57BL/6J; TISSUE=Lung;
 RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
 RA Carninci P., Shibata Y., Hayatsu N., Suganara Y., Shibata K., Itoh M.,
 RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RT "Normalisation and subcloning of cap-trapper-selected cDNAs to
 RT prepare full-length cDNA libraries for rapid discovery of new genes."; Genom. Res. 10:1617-1630(2000).
 RL [5]
 RN SEQUENCE FROM N.A.
 RP STRAIN=C57BL/6J; TISSUE=Lung;
 RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
 RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
 RA Kono H., Akiyama J., Nishi K., Kikunishi T., Tashiro H., Itoh M.,
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
 RA Yamamoto R., Matsunoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
 RA Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Watanabe M.,
 RA Takeda Y., Ishikawa T., Ozawa K., Tanaka T., Matsunaga S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
 RT "RIKEN integrated sequence analysis (RISA) system-384-format
 RT sequencing pipeline with 384 multicapillary sequencer.";

RL Genom. Res. 10:1757-1771(2000).
 RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Lung;
 RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
 RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
 RA Hayashida K., Hayatsu N., Hiramoto K., Hirooka T., Hirozane T.,
 RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
 RA Katoh H., Kawai J., Kojima Y., Kondo S., Kono H., Koyama S.,
 RA Koshimizu C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
 RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohnishi N., Okazaki Y.,
 RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
 RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
 RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
 RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
 RL Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AK087322; BAC39845.1; -
 DR MGD; MGI:104754; Snf1k.
 DR GO; GO:0016301; F.kinase activity; IEA.
 KW Kinase.
 FT NON TER
 SQ SEQUENCE 463 AA; 48966 MW; DA6327DB4AF27C78 CRC64;
 Query Match 38.1%; Score 48; DB 2; Length 463;
 Best Local Similarity 52.9%; Pred. No. 98;
 Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
 QY 5 SYNPAVDLYVFERK 21
 Db 14 SYNPAVDLYVFERK 30
 RESULT 11
 Q6ZNL8 PRELIMINARY; PRT; 764 AA.
 AC Q6ZNL8;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE FJ00263 protein (fragment).
 GN Name=FJ00263;
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 ON NCB1_TaxID=9606;
 RX SEQUENCE FROM N.A.
 RC TISSUE=Spleen;
 RA Jikuya H., Takano J., Kikuno R., Nagase T., Ohara O.;
 RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
 CC -i- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
 DR EMBL: AK131076; BAC85126.1; -
 DR HSSP; O63450; 1A06.
 DR GO; GO:0005524; P.ATP binding; IEA.
 DR GO; GO:0004674; P.protein serine/threonine kinase activity; IEA.
 DR GO; GO:0004713; P.protein-tyrosine kinase activity; IEA.
 DR GO; GO:0016740; P.transferase activity; IEA.
 DR GO; GO:0006468; P.protein amino acid phosphorylation; IEA.
 DR InterPro; IPR011009; Kinase like.
 DR InterPro; IPR000719; Ser kinase.
 DR InterPro; IPR002290; Ser Thr kinase.
 DR InterPro; IPR008271; Ser Thr kinase.
 DR InterPro; IPR001245; Tyr_kinase.
 DR InterPro; IPR000449; UBA.
 DR Pfam; PF00069; Pkinase; 1.
 DR ProDom; PD000001; Prc_kinase; 1.
 DR SMART; SM00220; S_TKc; 1.
 DR SMART; SM00219; TyrKc; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE; PS50030; UBA; 1.
 KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.


```

KW  ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ  SEQUENCE 783 AA; 84929 MW; 0A1B6067DD2EC90E CAC64;
OY  5 SYDNFALVDYLVFERTK 21
Db  330 SYNHPFALYLLERLK 346

Query Match 38.1%; Score 48; DB 2; Length 783;
Best Local Similarity 52.9%; Pred. No. 1.7e+02;
Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

RESULT 14
ID SN1L_HUMAN STANDARD; PRT; 786 AA.
AC P57059;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Probable serine/threonine-protein kinase SNF1LK (EC 2.7.1.37).
GN Name=SNF1LK;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RV
RP SEQUENCE FROM N.A.
RX MEDLINE=20289799; PubMed=10830953; DOI=10.1038/35012518;
RA Hattori M., Fujiyama A., Taylor T.D., Matanabe H., Yada T.,
RA Park H.-S., Toyoda A., Ishii K., Torki Y., Choi D.-K., Soeda E.,
RA Oki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K., Polley A.,
RA Menzel U., Delabre J., Kumpf K., Lehmann R., Patterson D.,
RA Reichwald K., Rump A., Schillnabel M., Schudy A., Zimmermann W.,
RA Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.,
RA Shintani A., Sasaki T., Nagamine K., Mitsuyma S., Antonarakis S.E.,
RA Mitoshima S., Shimizu N., Nordstieck G., Hornischer K., Brandt P.,
RA Scharfe M., Schoen O., Desario A., Reichelt J., Kauer G., Bloecker H.,
RA Ramser J., Beck A., Klages S., Hennig S., Rieselmann L., Degand E.,
RA Wehrmeyer S., Borzym K., Gardner K., Nizetic D., Francis F.,
RA Lehnach H., Reinhardt R., Yaspo M.-L.;
RT "The DNA sequence of human chromosome 21.";
RL Nature 405:311-319(2000).
CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family. SNF1
CC subfamily.
CC -!- SIMILARITY: Contains 1 UBA domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AP001751; BAA95536.1; -
DR HSSP; O14757; INVR.
DR Genew; HGNC:11142; SNF1LK.
DR MIM; 605705; -
DR InterPro; IPR011009; Kinase like.
DR InterPro; IPR000719; Prot. Kinase.
DR InterPro; IPR002290; Ser thr kinase.
DR InterPro; IPR008271; Ser thr_kin_AS.
DR InterPro; IPR000449; UBA-
DR Pfam; PF00069; Kinase; 1.
DR Prodom; PD000001; Prot_Kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00165; UBA; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00011; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS00030; UBA; 1.
KW ATP-binding; Serine/threonine-protein kinase; Transferase.

```

FT	DOMAIN	27	281	Protein kinase.
FT	NP_BIND	306	346	UBA
FT	NP_BIND	33	41	ATP (By similarity).
FT	BINDING	56	56	ATP (By similarity).
FT	ACT_SITE	149	149	Proton acceptor (By similarity).
FT	SEQUENCE	786 AA;	85252 MW;	FB44EBE7CAFB7CB1A CRC64;
SO	SEQUENCE	786 AA;	85252 MW;	FB44EBE7CAFB7CB1A CRC64;
Query Match	Similarity	38.1%;	Score 48;	DB 1;
Best Local	Similarity	52.9%;	Pred. No. 1.7e+02;	Length 786;
Matches	9;	Conservative	4;	Mismatches
Oy	5	SYDNFALVDYLVFERTK	21	
	:: : :			
Db	333	SYNHPAAYTYLLERLK	349	
Query Match	Similarity	38.1%;	Score 48;	DB 1;
Best Local	Similarity	52.9%;	Pred. No. 1.7e+02;	Length 786;
Matches	9;	Conservative	4;	Mismatches
Oy	5	SYDNFALVDYLVFERTK	21	
	:: : :			
Db	329	SYNHPAAYTYLLERLK	345	
Query Match	Similarity	38.1%;	Score 48;	DB 2;
Best Local	Similarity	52.9%;	Pred. No. 1.7e+02;	Length 798;
Matches	9;	Conservative	4;	Mismatches
Oy	5	SYDNFALVDYLVFERTK	21	
	:: : :			
Db	329	SYNHPAAYTYLLERLK	345	
Query Match	Similarity	38.1%;	Score 48;	DB 2;
Best Local	Similarity	52.9%;	Pred. No. 1.7e+02;	Length 798;
Matches	9;	Conservative	4;	Mismatches
Oy	5	SYDNFALVDYLVFERTK	21	
	:: : :			
Db	329	SYNHPAAYTYLLERLK	345	
Query Match	Similarity	38.1%;	Score 48;	DB 2;
Best Local	Similarity	52.9%;	Pred. No. 1.7e+02;	Length 798;
Matches	9;	Conservative	4;	Mismatches
Oy	5	SYDNFALVDYLVFERTK	21	
	:: : :			
Db	329	SYNHPAAYTYLLERLK	345	
Query Match	Similarity	38.1%;	Score 48;	DB 2;
Best Local	Similarity	52.9%;	Pred. No. 1.7e+02;	Length 798;
Matches	9;	Conservative	4;	Mismatches
Oy	5	SYDNFALVDYLVFERTK	21	
	:: : :			
Db	329	SYNHPAAYTYLLERLK	345	
Query Match	Similarity	38.1%;	Score 48;	DB 2;
Best Local	Similarity	52.9%;	Pred. No. 1.7e+02;	Length 798;
Matches	9;	Conservative	4;	Mismatches
Oy	5	SYDNFALVDYLVFERTK	21	
	:: : :			
Db	329	SYNHPAAYTYLLERLK	345	
Query Match	Similarity	38.1%;	Score 48;	DB 2;
Best Local	Similarity	52.9%;	Pred. No. 1.7e+02;	Length 798;
Matches	9;	Conservative	4;	Mismatches
Oy	5	SYDNFALVDYLVFERTK	21	
	:: : :			
Db	329	SYNHPAAYTYLLERLK	345	
Query Match	Similarity	38.1%;	Score 48;	DB 2;
Best Local	Similarity	52.9%;	Pred. No. 1.7e+02;	Length 798;
Matches	9;	Conservative	4;	Mismatches
Oy	5	SYDNFALVDYLVFERTK	21	
	:: : :			
Db	329	SYNHPAAYTYLLERLK	345	
Query Match	Similarity	38.1%;	Score 48;	DB 2;
Best Local	Similarity	52.9%;	Pred. No. 1.7e+02;	Length 798;
Matches	9;	Conservative	4;	Mismatches
Oy	5	SYDNFALVDYLVFERTK	21	
	:: : :			
Db	329	SYNHPAAYTYLLERLK	345	
Query Match	Similarity	38.1%;	Score 48;	DB 2;
Best Local	Similarity	52.9%;	Pred. No. 1.7e+02;	Length 798;
Matches	9;	Conservative	4;	Mismatches
Oy	5	SYDNFALVDYLVFERTK	21	
	:: : :			
Db	329	SYNHPAAYTYLLERLK	345	
Query Match	Similarity	38.1%;	Score 48;	DB 2;
Best Local	Similarity	52.9%;	Pred. No. 1.7e+02;	Length 798;
Matches	9;	Conservative	4;	Mismatches
Oy	5	SYDNFALVDYLVFERTK	21	
	:: : :			
Db	329	SYNHPAAYTYLLERLK	345	
Query Match	Similarity	38.1%;	Score 48;	DB 2;
Best Local	Similarity	52.9%;	Pred. No. 1.7e+02;	Length 798;
Matches	9;	Conservative	4;	Mismatches
Oy	5	SYDNFALVDYLVFERTK	21	
	:: : :			
Db	329	SYNHPAAYTYLLERLK	345	
Query Match	Similarity	38.1%;	Score 48;	DB 2;
Best Local	Similarity	52.9%;	Pred. No. 1.7e+02;	Length 798;
Matches	9;	Conservative	4;	Mismatches
Oy	5	SYDNFALVDYLVFERTK	21	
	:: : :			
Db	329	SYNHPAAYTYLLERLK	345	
Query Match	Similarity	38.1%;	Score 48;	DB 2;
Best Local	Similarity	52.9%;	Pred. No. 1.7e+02;	Length 798;
Matches	9;	Conservative	4;	Mismatches
Oy	5	SYDNFALVDYLVFERTK	21	
	:: : :			

```

DT 01-DEC-2001 (TEMBLrel. 19, last sequence update)
DT 25-OCT-2004 (TEMBLrel. 28, last annotation update)
DE Putative pms collagenase.
GN Name=P0487H02.27; Synonyms=P0682B08.8;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophytes; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Eriharcoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=12447438; DOI=10.1038/nature01184;
RA Sasaki T., Matsumoto T., Yamamoto K., Sakata K., Baba T., Katayose Y.,
RA Wu J., Nilmura Y., Cheng Z., Nagamura Y., Antonio B.A., Kanamori H.,
RA Hosokawa S., Maekawa M., Arikawa K., Chiden Y., Hayashi M.,
RA Okamoto M., Ando T., Aoki H., Arita K., Hamada M., Harada C.,
RA Hijishita S., Honda M., Ichikawa Y., Idenuma A., Iijima M., Ikeda M.,
RA Ikono M., Itoh S., Itoh T., Itoh Y., Itoh Y., Iwabuchi A., Kamiya K.,
RA Karaeawa W., Katagiri S., Kikuta A., Kobayashi N., Kono I.,
RA Machita K., Maehara T., Mizuno H., Mizubayashi T., Mukai Y.,
RA Nagasaki H., Negishi M., Nakama Y., Nakamichi T., Nakamura M.,
RA Namiki N., Negishi M., Ohta I., Ono N., Saji S., Sakai K., Shibata M.,
RA Shimokawa T., Shomura A., Song J., Takazaki Y., Terasawa K., Tsuji K.,
RA Waki K., Yamagata H., Yamane H., Yoshiki S., Yoshihara R., Yukawa K.,
RA Zhang H., Iwama H., Endo T., Ito H., Hann J.H., Kim H.I., Eun M.Y.,
RA Yano M., Jiang J., Gojobori T.;
RT "The genome sequence and structure of rice chromosome 1.";
RL Nature 420:312-316(2002).
EMBL: AP002883; BAB67867.1; -.
DR EMBL: AP003578; BAB60938.1; -.
DR Gramene; Q93M04; -.
DR GO; GO:0009279; C:outer membrane (sensu Gram-negative Bacteria); IEA.
DR InterPro; IPR008969; Carboxypep_reg.
DR InterPro; IPR008454; Cna_B.
DR InterPro; IPR001809; Outsurface.
DR Pfam; PF05738; Cna_B; 2.
KM Collagen.
SQ SEQUENCE 1193 AA; 130639 MW; 6BAE0A9E1BCE0D5 CRC64;

Query Match 38.1%; Score 48; DB 2; Length 1193;
Best Local Similarity 42.9%; Pred. No. 2.6e+02;
Matches 9; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

QY 2 ISVSYNPALVDYLVFERTKS 22
Db 1014 IDVGQDDISGIDFVFERPES 1034

RESULT 17
Q6FS04 PRELIMINARY; PRT; 232 AA.
ID Q6FS04
AC Q6FS04;
DT 05-JUL-2004 (TEMBLrel. 27, Created)
DT 05-JUL-2004 (TEMBLrel. 27, Last sequence update)
DE Similar to sp|P48353 Saccharomyces cerevisiae YMR161w HLJ1.
GN ORFNames=CAG10H04999;
OS Candida glabrata CBS138.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=284553;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=CBS138;
RG Genolevures;
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA Lafontaine I., de Montigny J., Marck C., Neuveglise C., Talla E.,
RA Goffard N., Franquel L., Aigle M., Anthouard V., Babour A., Babbe V.,
RA Barnay S., Blanchin J., Beckerich J.M., Beyne E., Bleykasten C.,
RA Boisrame A., Boyer J., Cattoic L., Confalonieri F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantreay F., Hennequin C., Janniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Koszul K., Lemaire M., Lesur I., Ma L., Muller H.,

```

```

RA Nicaud J.M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,
RA Pallen S., Porter S., Richard G.F., Straub M.L., Sileau A.,
RA Swenne D., Tekala F., Wesolowski-Louvel M., Weathof E., Wirth B.,
RA Zanotti-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gallardin C., Weissenbach J.,
RA Wincker P., Souciet J.L.;
RT "Genome evolution in yeasts.";
RL Nature 430:35-44(2004).
EMBL: CR380954; CAG59923.1; -.
DR GO; GO:0051082; F:unfolded protein binding; IEA.
DR GO; GO:0006457; P:protein folding; IEA.
DR InterPro; IPR001623; DnaU_N.
DR InterPro; IPR003095; Hsp_DnaU.
DR Pfam; PF00226; DnaU; 1.
DR PRINTS; PR00625; DNAJPROTEIN.
DR SMART; SM00271; DnaJ; 1.
DR PROSITE; PS00636; DnaJ_1; 1.
DR PROSITE; PS50076; DnaJ_2; 1.
SQ SEQUENCE 232 AA; 26494 MW; 7F4ED03FB3E44D7 CRC64;

```

```

Query Match 37.3%; Score 47; DB 2; Length 232;
Best Local Similarity 45.5%; Pred. No. 68;
Matches 10; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 4 VSYDNFALVDYLVFERTKSDTD 25
Db 15 LSHDKAFYDILNVERSSDSD 36

RESULT 18
Q97AL5 PRELIMINARY; PRT; 321 AA.
ID Q97AL5
AC Q97AL5;
DT 01-OCT-2001 (TEMBLrel. 18, Created)
DT 01-OCT-2001 (TEMBLrel. 18, Last sequence update)
DT 01-MAR-2004 (TEMBLrel. 26, Last annotation update)
DE Hypothetical protein TVG079595.
GN Name=TVG079595; OrderedLocustNames=TV0795;
OS Thermoplasma volcanium.
OC Archaea; Euryarchaeota; Thermoplasmatia; Thermoplasmatiales;
OC Thermoplasmales; Thermoplasma.
OX NCBI_TaxID=50339;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=GSS1 / DSM 4299 / JCM 9571;
RX MEDLINE=20570466; PubMed=11121031; DOI=10.1073/pnas.97.26.14257;
RA Kawashima T., Amano N., Koike H., Makino S.-I., Higuchi S.,
RA Kawashima-Ohya Y., Watanabe K., Yamazaki M., Kanehori K., Kawamoto T.,
RA Nunoshiba T., Yamamoto Y., Aramaki H., Makino K., Suzuki M.;
RT "Archaeal adaptation to higher temperatures revealed by genomic
RT sequence of Thermoplasma volcanium.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:14257-14262(2000).
EMBL: AP000993; BAB59937.1; -.
DR HSP; 074023; IMGT.
DR GO; GO:0004519; F:endonuclease activity; IEA.
DR GO; GO:000308; F:methylated-DNA-[protein]-cysteine S-methylt...; IEA.
DR GO; GO:0006281; P:DNA repair; IEA.
DR InterPro; IPR007581; Endonuc_V.
DR InterPro; IPR001497; Methyltransf_1.
DR Pfam; PF04493; Endonuclease_5; 1.
DR Pfam; PF01035; Methyltransf_1; 1.
DR TIGRFAMs; TIGR00589; ogt; 1.
KM Complete proteome; Hypothetical protein.
SQ SEQUENCE 321 AA; 36049 MW; FC3220C1E37B7ACD CRC64;

```

```

Query Match 37.3%; Score 47; DB 2; Length 321;
Best Local Similarity 57.1%; Pred. No. 95;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 ISVSYNPALVDYLV 15
Db 2 LSVYDNFALVDYLV 15

```

RESULT 19
ID 005583 PRELIMINARY; PRT: 330 AA.
AC 005583;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE V4R267CP
GN ORFNames=YDR267C;
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RA Le T.;
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z.,
RA Favello A., Fulton L., Gattung S., Greco T., Kirsten J., Kucaba T.,
RA Hallsworth K., Hawkins J., Hillier L., Jier M., Johnson D.,
RA Johnston L., Langston Y., Latreille P., Le T., Mardis E., Meneses S.,
RA Miller N., Nhan M., Pauley A., Peluso D., Rifken L., Riles L.,
RA Titch N., Tveaekis E., Vignati D., Wilcox L., Wohlman P., Vaudin M.,
RA Wilson R., Waterston R.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RA Jia Y., Cherry J.M.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 7 WD repeats.
DR EMBL; U51030; AAB64456.1; -
DR PIR; S70127; S70127.
DR HSSP; P16649; 1BRJ.
DR IntAct; Q05583; -
DR SGD; S000002675; YDR267C.
DR GO; GO:0005737; C:cytoplasm; IDA.
DR GO; GO:0005634; C:nucleus; IDA.
DR InterPro; IPR011569; His bind.
DR InterPro; IPR001680; WD40.
DR InterPro; IPR011046; WD40_like.
DR Pfam; PF00400; WD40_7.
DR PRINTS; PR00320; GPROTEINRPT.
DR PRODOM; PD152455; His bind; 2.
DR PRODOM; PD000018; WD40_2.
DR SMART; SM00320; WD40_7.
DR PROSITE; PS00678; WD_REPEATS_1; UNKNOWN_1.
DR PROSITE; PS50082; WD_REPEATS_2; 5.
DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
KM Repeat; WD repeat.
SQ SEQUENCE 330 AA; 37274 MW; B4CCAA3125FD666 CRC64;

DE KIAA0781 protein (Fragment).
GN Name=KIAA0781;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Takemori H.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR EMBL; AB084424; BAB91442.1; -
DR HSSP; P49137; INY3.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR011009; Kinase like.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR002290; Ser_Thr_kinase.
DR InterPro; IPR008271; Ser_Thr_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR InterPro; IPR000449; UBA.
DR Pfam; PF00069; Pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR PRODOM; PD000001; Prot kinase; 1.
DR SMART; SM00220; S_TKC_1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS50030; UBA; 1.
KM ATP-binding; kinase; Serine/threonine-protein kinase; Transferase.
FT NON_TER 1 1
FT NON_TER 346 346
SQ SEQUENCE 346 AA; 40073 MW; BD883EA07F612EE3 CRC64;

Query Match 37.3%; Score 47; DB 2; Length 346;
Best Local Similarity 50.0%; Pred. No. 1e+02;
Matches 9; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 5 SYNFPALVDLVPERKTS 22
DB 320 SYNHPAIVFLVRLKLS 337

RESULT 21
ID 09BJZ1 PRELIMINARY; PRT: 505 AA.
AC 09BJZ1;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Beta-1,4-endoglucanase (EC 3.2.1.4) (Fragment).
GN Name=eng-1a;
OS Meloidogyne incognita (southern root-knot nematode).
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;
OC Tylenchoidea; Heteroderidae; Meloidogyninae; Meloidogyne.
OX NCBI_TaxID=6306;
RN [1]
RP SEQUENCE FROM N.A.
RA Ledger T.N., Jaubert S., Carot J., Arnaud L., Abad P., Rosso M.N.;
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF323087; AKK21882.1; -
DR HSSP; P07103; IEG2.
DR GO; GO:0008810; F:cellulase activity; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro; IPR001919; Bac celose-bind.
DR InterPro; IPR008965; Cellul bind.
DR InterPro; IPR001547; Glyco_hydro_5.
DR Pfam; PF00150; Cellulase; 1.
DR SMART; SM00637; CBD_II; 1.
DR PROSITE; PS00659; GLYCOSYL_HYDROL_F5; UNKNOWN_1.

KW Glycosidase; Hydrolase.
 FT NON TER 1
 SQ SEQUENCE 505 AA; 53302 MW; C4E4551D70C7F9E CRC64;

Query Match 37.3%; Score 47; DB 2; Length 505;
 Best Local Similarity 43.5%; Pred. No. 1.5e+02;
 Matches 10; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

Qy 3 VSVDNFPALVDVLPFERTKSDTD 25
 Db 459 NVTGNMFLPDYVITFAGKQYTD 481

RESULT 22
 MALF PHOLL STANDARD; PRT; 518 AA.
 AC 07N964;
 DT 25-OCT-2004 (Rel. 45, Created)
 DT 25-OCT-2004 (Rel. 45, Last sequence update)
 DE Maltose transport system permease protein malF.
 GN Name=malF; OrderedLocNames=plu0459;
 OS Photobacterium luminescens (subsp. laumondii).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Photobacterius.
 OX NCBI_TaxID=141679;
 RN SEQUENCE FROM N.A.
 RC STRAIN=TT01;
 RX MEDLINE=22957627; PubMed=14528314; DOI=10.1038/nbt886;
 RA Duchaud E., Ruelink C., Frangeul L., Buchrieser C., Givaudan A.,
 TAouit S., Bocs S., Bouraux-Ende C., Chantier M., Charles J.-F.,
 Dassa E., Derose R., Derzelle S., Freysinet G., Gaudriault S.,
 RA Medigue C., Lanois A., Powell K., Siguler P., Vincent R., Wingate V.,
 RA Zouine M., Glaeser P., Boemare N., Danchin A., Kunst F.,
 RT "The genome sequence of the entomopathogenic bacterium Photobacterium
 luminescens.";
 RL Nat. Biotechnol. 21:1307-1313(2003).
 CC -1- FUNCTION: Part of a binding-protein-dependent transport system.
 CC Probably responsible for the translocation of the substrate across
 CC the membrane (By similarity).
 CC -1- SUBUNIT: The complex is composed of two ATP-binding proteins
 CC (malK), two transmembrane proteins (malG and malF) and a solute-
 CC binding protein (malE) (By similarity).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
 CC (By similarity).
 CC -1- SIMILARITY: Belongs to the binding-protein-dependent transport
 CC system permease family. MalG subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: BX571860; CAE12754.1; -
 CC Photolier; plu0459; -
 DR InterPro: IPR000515; BPD_transp.
 DR Pfam: PF00528; BPD_transp_1; 1.
 DR PROSITE: PS50928; ABC_TMI; 1.
 KW Complete proteome; Inner membrane; Sugar transport; Transmembrane;
 KM Transport.
 FT TRANSMEM 15 34 potential.
 FT TRANSMEM 41 58 potential.
 FT TRANSMEM 73 95 potential.
 FT TRANSMEM 281 303 potential.
 FT TRANSMEM 318 340 potential.
 FT TRANSMEM 377 399 potential.
 FT TRANSMEM 488 510 potential.
 SQ SEQUENCE 518 AA; 57617 MW; 50BA4AD4EBB63C296 CRC64;

Query Match 37.3%; Score 47; DB 1; Length 518;
 Best Local Similarity 33.3%; Pred. No. 1.6e+02;
 Matches 7; Conservative 9; Mismatches 5; Indels 0; Gaps 0;

Qy 2 ISVSDNFPALVDVLPFERTKS 22
 Db 88 IAIATVSYISNQLTFERRAS 108.

RESULT 23
 06D2L3 PRELIMINARY; PRT; 527 AA.
 ID 06D2L3;
 AC 06D2L3;
 DT 25-OCT-2004 (TREMBLrel. 28, Created)
 DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
 DE Maltate:quinone oxidoreductase (EC 1.1.99.16).
 GN Name=mgc; OrderedLocNames=ECA3082;
 OS Erwinia carotovora (subsp. atroseptica) (pectobacterium atrosepticum).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Pectobacterium.
 OX NCBI_TaxID=29471;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SCRI 1043 / ATCC BAA-672;
 RX PubMed=15263089; DOI=10.1073/pnas.0402424101;
 RA Bell K.S., Sebatia M., Pritchard L., Holden M.T.G., Hyman L.J.,
 RA Holeva M.C., Thomson N.R., Bentley S.D., Churcher L.J.C., Mungall K.,
 RA Atkin R., Bason N., Brooks K., Chillingworth T., Clark K., Doggett J.,
 RA Fraser A., Hance Z., Hauser H., Jagers K., Moule S., Norbertczak H.,
 RA Ormond D.P., Price C., Quail M.A., Sanders M., Walker D., Whitehead S.,
 RA Salmond G.P.C., Birch P.R.J., Parkhill J., Toth I.K.,
 RT "Genome sequence of the enterobacterial phytopathogen Erwinia
 RT carotovora subsp. atroseptica and characterization of virulence
 RT factors.";
 RL Proc. Natl. Acad. Sci. U.S.A. 101:11105-11110(2004).
 RU EMBL: BX50851; CAG75981.1; -
 DR GO: GO:0008924; F:maltate dehydrogenase (acceptor) activity; IEA.
 DR GO: GO:0016491; F:oxidoreductase activity; IEA.
 DR GO: GO:0006099; F:tricarboxylic acid cycle; IEA.
 DR InterPro: IPR006231; mal_quin_oxido.
 DR Pfam: PF06039; Mgc; 1.
 DR TIGRFAMs: TIGR01320; mal_quin_oxido; 1.
 KW Complete proteome; Oxidoreductase.
 SQ SEQUENCE 527 AA; 58123 MW; 35D674A15C000ACD CRC64;

Query Match 37.3%; Score 47; DB 2; Length 527;
 Best Local Similarity 50.0%; Pred. No. 1.6e+02;
 Matches 11; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

Qy 4 VSVDNFPALVDVLPFERTKSDTD 25
 Db 376 VGIDNFDVLYVLSQWMDDD 397

RESULT 24
 088S13 PRELIMINARY; PRT; 634 AA.
 ID 088S13;
 AC 088S13;
 DT 01-JUN-2003 (TREMBLrel. 24, Created)
 DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
 DE Cadmium transporting P-type ATPase.
 GN Name=cadA; OrderedLocNames=lp_345;
 OS Lactobacillus plantarum.
 OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
 OC Lactobacillus.
 OX NCBI_TaxID=1590;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NCIM 8826 / WCF51;
 RX MEDLINE=22480296; PubMed=12566566; DOI=10.1073/pnas.0337704100;
 RA Kleerebezem M., Boekhorst J., van Kranenburg R., Molenaar D.,

RA Kuipers O.P., Leer R., Tarchini R., Peters S.A., Sandbrink H.M.,
 RA Fiers M.W.E.J., Stiekema W., Klein Lankhorst R.M., Bron P.A.,
 RA Hoffer M.M., Nierop Groot M.N., Kerkhoven R., De Vries M., Uuring B.,
 RA De Vos W.M., Stezen R.J.,
 RT "Complete genome sequence of *Lactobacillus plantarum* WCFS1.";
 RL Proc. Natl. Acad. Sci. U.S.A. 100:1990-1995 (2003).
 DR EMBL: AL935262; CAD5533.1; -;
 DR GO: GO:0016021; C: integral to membrane; IEA.
 DR GO: GO:0005524; F: ATP binding; IEA.
 DR GO: GO:0015662; F: ATPase activity, coupled to transmembrane m...; IEA.
 DR GO: GO:0016787; F: hydrolase activity; IEA.
 DR GO: GO:0016818; F: hydrolase activity, acting on acid anhydrid...; IEA.
 DR GO: GO:0016820; F: hydrolase activity, acting on acid anhydrid...; IEA.
 DR GO: GO:0046872; F: metal ion binding; IEA.
 DR GO: GO:0046873; F: metal ion transporter activity; IEA.
 DR GO: GO:0008152; F: metabolism; IEA.
 DR GO: GO:0030001; F: metal ion transport; IEA.
 DR InterPro: IPR006416; ATPase-IB_hvy.
 DR InterPro: IPR001757; ATPase-IB_E2.
 DR InterPro: IPR005834; Dehal_Like_hydro.
 DR InterPro: IPR008250; E1-E2_ATPase_reg.
 DR InterPro: IPR006404; Heavy_mec_ATPase.
 DR Pfam: PF00122; E1-E2_ATPase; 1.
 DR Pfam: PF00702; Hydrolase; 1.
 DR PRINTS: PR00119; CATRAPASE.
 DR TIGRFAMs: TIGR01512; ATPase-IB2_Cd; 1.
 DR TIGRFAMs: TIGR01525; ATPase-IB_hvy; 1.
 DR TIGRFAMs: TIGR01494; ATPase_P_Type; 2.
 DR PROSITE: PS00154; ATPASE_E1_E2; UNKNOWN_1.
 DR PROSITE: PS01229; COF_2; UNKNOWN_1.
 KW Complete proteome.
 SQ SEQUENCE 634 AA; 67508 MW; C20120EB81508287 CRC64;

Query Match 37.3%; Score 47; DB 2; Length 634;
 Best Local Similarity 52.6%; Pred. No. 1.9e+02;
 Matches 10; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 7 DNPAALVDYLVFERTKSDTD 25
 DB 301 DTFPAKVDTLVFDKGTGLTE 319

RESULT 25
 ID Q6AZE2 PRELIMINARY; PRT; 711 AA.
 AC Q6AZE2;
 DT 25-OCT-2004 (T-EMBLrel. 28, Created)
 DT 25-OCT-2004 (T-EMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (T-EMBLrel. 28, Last annotation update)
 DE SIK2 protein (Fragment).
 GN Name=SIK2;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RX PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altshuler S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heish F.,
 RA Diatchenko L., Marulita K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Tsohlyuki S., Carninci P., Prange C.,
 RA Baha S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mulhaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richard S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzyzanski M.I., Skalek U., Smallos D.E., Schnerch A., Schein J.E.,
 RA Jones S.V., Watra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RA Director MGC Project;
 RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC078150; AAH78150.1; -;
 DR GO: GO:0005524; F: ATP binding; IEA.
 DR GO: GO:0004672; F: protein kinase activity; IEA.
 DR GO: GO:0004686; F: protein amino acid phosphorylation; IEA.
 DR InterPro: IPR001009; Kinase_Like.
 DR InterPro: IPR000719; Prot_Kinase.
 DR Pfam: PF00069; Kinase; 1.
 DR ProDom: PD000001; Prot_kinase; 1.
 DR PROSITE: PS50030; UBA; 1.
 FT NON TER 1
 SQ SEQUENCE 711 AA; 79410 MW; 2D3BA33FEB21CC21 CRC64;

Query Match 37.3%; Score 47; DB 2; Length 711;
 Best Local Similarity 50.0%; Pred. No. 2.2e+02;
 Matches 9; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 5 SYDNPAALVDYLVFERTKS 22
 DB 107 SYNHPAALYFLVLRKLS 124

Search completed: May 9, 2005, 12:27:29
 Job time: 175 secs

This Page Blank (usp10)